Portfolio Optimization in a Jump-Diffusion Market with Durability and Local Substitution: A Penalty Approximation of a Singular Control Problem

by

Nina Holden

Thesis for the degree of

Master of Science

(Master i Anvendt matematikk og mekanikk)

Faculty of Mathematics and Natural Sciences
University of Oslo

November 2010

Contents

Abstract

The main purpose of this thesis is to study a singular finite-horizon portfolio optimization problem, and to construct a penalty approximation and numerical scheme for the corresponding Hamilton-Jacobi-Bellman (HJB) equation. The driving stochastic process of the portfolio optimization problem is a Lévy process, and the HJB equation of the problem is a non-linear second order degenerate integro-partial differential equation subject to gradient and state constraints. We characterize the value function of the optimization problem as the unique constrained viscosity solution of the HJB equation.

Our penalty approximation is obtained by studying a non-singular version of the original optimization problem. The original HJB equation is difficult to solve numerically because of the gradient constraint, and in the penalty approximation the gradient constraint is incorporated in the HJB equation as a penalty term. We prove that the solution of the penalty approximation converges to the solution of the original problem. We also construct a numerical scheme for the penalty approximation, prove convergence of the numerical scheme, and present results from numerical simulations.

Other new results of the thesis include some explicit solution formulas for our finite-horizon optimization problem, new constraints on the explicit time-independent solution formulas found in [8], and a lemma concerning the characterization of viscosity solutions by use of so-called subdifferentials and superdifferentials.

Chapter 1
Introduction

We will study a problem of optimal consumption and portfolio selection. An investor has the opportunity to invest her wealth in an uncertain asset and a certain asset, and can also consume wealth. Her goal is to maximize expected utility, where utility is a function of consumption and wealth.

The optimization problem can be formulated as a non-linear integro-PDE with gradient constraint, and the objective of this thesis is to describe properties of its solution, and construct and analyze approximation methods for the PDE. The original PDE is difficult to solve numerically, and therefore we will approximate the PDE by so-called penalty approximations.

Some features of the optimization problem compared to the standard Merton problem, is that

(1) the optimization problem captures the notions of durability and local substitution, i.e., the agent's satisfaction at a certain time depends on both current and past consumption, and her satisfaction is not changed much if the time for consumption is changed slightly, and

(2) the logreturns of the uncertain assets are not necessarily normally distributed, but are modelled as a Lévy process.

Portfolio optimization problems satisfying (1) were first studied extensively by Hindy and Huang in [31] for a marked modelled by geometric Brownian motion. The main purpose of introducing (1), is that it corresponds well with realistic assumption about the agent's preferences.

In [8] and [9] Benth, Karlsen and Reikvam extend Hindy and Huang's model to Lévy processes. Empirical work by Eberlein and Keller [21] and Rydberg [39] show that the normal distribution fits the logreturn data poorly, and several papers ([6], [21]) suggest to generalize the distribution. The Lévy process is an appropriate generalization that fits the empirical data well.

The problem studied in [8] and [9] is an infinite-horizon version of our problem. It is shown in these articles that the value function is the unique constrained viscosity solution of the associated Hamilton-Jacobi-Bellman (HJB) equation, which is a non-linear integro-PDE subject to a gradient constraint. The first goal of this thesis is to adapt their results to a finite time-horizon. Some explicit solution formulas are found in [8], and some similar explicit solution formulas for the finite-horizon case will be found here. We will also formulate a new constraint on the parameters for one of the explicit solution formulas found in [8].

The gradient constraint of the HJB equation, which comes from the incorporated effects of the notions of durability and local substitution, makes the integro-PDE very difficult to solve. This thesis presents a penalty approximation to the HJB equation, and it is proved that the solution of the penalty approximation converges to the solution of the original HJB equation. The penalty approximation is the HJB equation of a continuous version of the optimization problem, where only absolutely continuous controls with bounded derivatives are allowed. Four different proofs of convergence are presented: one proof only valid for strictly increasing stock prices, two proofs for the general case, and one proof that is based on a strong comparison principle. The strong comparison principle will not be proved, as the proof is harder than the proof of the weak comparison principle, but references to related proofs found in other articles will be given.

We will also construct a numerical scheme for the penalty approximation, and prove that the numerical scheme is monotone, stable and consistent. Monotonicity, stability and consistency imply that the numerical solution converges to the solution of the penalty approximation. The form of our numerical scheme is taken from [12], while the proof of convergence uses techniques and results from [5] and [15]. We will implement the scheme, study its rate of convergence and time complexity, and discuss various properties of the value function and the calculated optimal controls. We will also compare the performance of different boundary value conditions for the scheme. For a certain choice of utility functions, the dimension of the HJB equation can be reduced. We will construct a penalty approximation and a numerical scheme also for this problem, and study convergence properties of the scheme.

Immediately after this introduction the reader will be introduced to some of the theory the project is based on. The experienced reader may want to skip it, while others may read it as an introduction to new topics or use it for reference.

In Part I the stochastic optimization problem will be stated, and the viscosity solution theory developed in [8], [9] for the finite-horizon case is adapted to our case. Notation and assumptions made on functions and constants involved are given in Chapter 3. Part II also contains a heuristic derivation of the HJB equation, some properties of the value function are proved, and some explicit solution formulas are found.

Part II contains the main new result developed in this thesis. The penalty approximation is derived, it is proved that the penalization problem is well-posed within the theory of viscosity solutions, and it is proved that the solution of the penalty approximation converges uniformly to the solution of the original problem on compact subsets.

In Part III the numerical scheme that solves the penalty problem is presented, and numerical results are presented and discussed.

None of the proofs given in this thesis are direct copies of proofs found other places, but many of the proofs are inspired by proofs of similar theorems found in articles, book, etc. All nontrivial proofs that are based on proofs from other sources will give a reference to the source. If many proofs in a section or chapter are inspired by the same article or book, the reference may be given at the beginning of the section/chapter.

Acknowledgments

Thanks to my supervisor Kenneth Hvistendahl Karlsen for providing me with an interesting problem, giving me good advices and support in the problem solving process, and always being available for discussions and questions. Thanks to my parents, Helge Holden and André Damslora for feedback on my drafts. Also thanks to my fellow maths students for the past 5 years at university and all good memories from these years.

Chapter 2
Theoretical background

This chapter gives a short introduction to some of the theory the thesis is based on. The purpose is partly to place the problem we consider in a greater context, partly to present results and introduce concepts that will be needed later. Note that the introduction is very short, and that it is focused mostly on theory directly relevant for other parts of the thesis. The reader is referred to the literature for further information and more thorough introductions.

2.1 Measure theory

Some basic measure theory is needed when working with stochastic processes and when proving convergence and integrability of functions. The definitions and theorems can be found in [14] and [24].

Definition 2.1 (σ-algebra). *A collection $\mathcal{E}$ of subsets of a set E is called a σ-algebra on E if the following conditions are satisfied:*

1. *The empty set $\emptyset \in \mathcal{E}$.*
2. *$\mathcal{E}$ is closed under complementation: If $A \in \mathcal{E}$, its complement $E \backslash A \in \mathcal{E}$.*
3. *$\mathcal{E}$ is closed under countable unions: If $A_1, A_2, \ldots \in \mathcal{E}$, then $A := A_1 \cup A_2 \cup \ldots \in \mathcal{E}$.*

We will mostly be working with subsets of $\mathbb{R}^n$ and the Borel σ-algebra on $\mathbb{R}^n$, defined by the following definition.

Definition 2.2 (Borel σ-algebra). *The smallest σ-algebra containing all open sets of a subset $E \subset \mathbb{R}^n$, is called the Borel σ-algebra, and it is denoted by $\mathcal{B}(E)$ or simply $\mathcal{B}$.*

The *measure* of an element in a σ-algebra, defined by the definition below, is often interpreted intuitively as the size or volume or the element.

Definition 2.3 (Measure). *Let $\mathcal{E}$ be a σ-algebra of subsets of E. Then $(E, \mathcal{E})$ is called a measurable space. A function $\mu : \mathcal{E} \to [0, \infty]$ is called a measure on $(E, \mathcal{E})$ if it satisfies the following conditions:*

1. *$\mu(\emptyset) = 0$.*
2. *For any sequence $\{A_n\}_{n\in\mathbb{N}}$ of disjoint sets such that $A_n \in \mathcal{E}$ for all $n \in \mathbb{N}$,*

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

An element $A \in \mathcal{E}$ is called a measurable set, and $\mu(A)$ is called its measure.

Many results we want to prove using measures, are only valid for so called *Radon measures*.

Definition 2.4 (Radon measure). *Let $E \subseteq \mathbb{R}$. A Radon measure on $(E, \mathcal{B})$ is a measure μ such that for every compact measurable set $B \in \mathcal{B}$, $\mu(B) < \infty$.*

We will be working with the *Lebesgue measure* in this thesis. This is the measure that defines $\mu(B)$ to be the n-dimensional volume of B for all balls B in $\mathbb{R}^n$. We see immediately that the Lebesgue measure is a Radon measure.

Many properties we will consider in measure theory are assumed to hold only so-called *almost everywhere.*

Definition 2.5 (Almost everywhere). *If some property holds everywhere on $\mathbb{R}$, except for a measurable set of Lebesgue measure zero, we say that the property holds almost everywhere, abbreviated "a.e.".*

In many applications we are interested in evaluating a measure on a set of the form $\{x \in E : f(x) \in A\}$, where f is a function between two measurable spaces $(E, \mathcal{E})$ and $(F, \mathcal{F})$, and $A \in \mathcal{F}$. This motivates the following definition.

Definition 2.6 (Measurable function). *Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be two measurable spaces. A function $f : E \to F$ is called measurable if, for any measurable set $A \in \mathcal{F}$, the set*

$$f^{-1}(A) = \{x \in E : f(x) \in A\}$$

is a measurable subset of E.

We want to define the integral of a measurable function by approximation by so-called simple functions. A simple function $f : [0, T] \to \mathbb{R}$ is a function that can be written on the form

$$f(t) = \sum_{i=1}^{\infty} \chi_{E_i}(t) u_i$$

for all $t \in [0, T]$, where E_i is a measurable set for $i = 1, 2, \ldots$ and $u_i \in \mathbb{R}$, and χ is the indicator function. If $f : \mathbb{R} \to [0, \infty)$ is a non-negative, measurable function it is possible, by an approximation of f with simple functions, to define the Lebesgue integral

$$\int_{\mathbb{R}} f \, dx.$$

The function f is said to be *integrable* if this integral is finite. We let f^+ and f^- denote, respectively, the positive and negative part of f for any function f, i.e., $f = f^+ - f^-$ for non-negative functions f^+ and f^-. If f is measurable, but not non-negative, we may define

$$\int_{\mathbb{R}} f \, dx = \int_{\mathbb{R}} f^+ \, dx - \int_{\mathbb{R}} f^- \, dx$$

if both terms on the right-hand side are finite. The definition of the integral given here agrees with the Riemann integral if f is Riemann integrable.

The following theorems are useful when proving that a function is integrable.

Theorem 2.7 (Monotone convergence theorem). *Assume the functions* $\{f_k\}_{k\in\mathbb{N}}$ *are measurable, with*

$$0 \leq f_1 \leq f_2 \leq \dots,$$

and that $\int_{\mathbb{R}} f_k\, dx$ *is bounded. Then*

$$\int_{\mathbb{R}} \lim_{k\to\infty} f_k\, dx = \lim_{k\to\infty} \int_{\mathbb{R}} f_k\, dx.$$

Theorem 2.8 (Dominated convergence theorem). *Assume the functions* $\{f_k\}_{k\in\mathbb{N}}$ *are integrable, that* $f_k \to f$ *a.e., and that* $|f_k| \leq g$ *a.e. for some integrable function* g. *Then* f *is integrable, and*

$$\int_{\mathbb{R}} f_k\, dx \to \int_{\mathbb{R}} f\, dx.$$

2.2 Stochastic analysis and Lévy processes

The driving stochastic process in the problem of this thesis, is the Lévy process. In this section the Lévy process will be defined, and some of its basic properties will be stated. First we need to give a short introduction to general stochastic processes. The presentation follows [14] closely.

Definition 2.9 (Probability space). $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space if* Ω *is a set of scenarios,* $\mathcal{F}$ *is a* σ*-algebra on* Ω*, and* $\mathbb{P}$ *is a positive finite measure on* $(\Omega, \mathcal{F})$ *with total mass 1.*

Definition 2.10 (Random variable). *A random variable* X *taking values in* A*, is a measurable function* $X : \Omega \to A$*, where* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space.*

We use E to denote the expected value of X.

Definition 2.11 (Convergence in probability). *A sequence* $\{X_n\}_{n\in\mathbb{N}}$ *of random variables on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is said to converge in probability to a random variable* X *if, for each* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$$

A Poisson random variable is an example of a random variable, and it will be used later when defining the Lévy process.

Definition 2.12 (Poisson random variable). *A Poisson random variable with parameter* λ *is an integer-valued random variable* X *such that*

$$\mathbb{P}(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}$$

for all $n \in \mathbb{N}$.

Now the concept of stochastic processes is introduced.

Definition 2.13 (Stochastic process). *A stochastic process is a family of random variables indexed by time.*

If the time parameter t is continuous, and the time interval is bounded, the stochastic process can be written as $(X_t)_{t\in[0,T]}$. For each realization $\omega \in \Omega$, the trajectory $X(\omega)$: $t \rightarrow X_t(\omega)$ defines a function of time called the *sample path* of the process.

As time goes on, more information about a stochastic process is revealed to the observer. We may formulate this mathematically by introducing a so-called *filtration* $(\mathcal{F}_T)_{t\in[0,T]}$.

Definition 2.14 (Filtration). *A filtration or* information flow *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *is an increasing family of* σ*-algebras* $(\mathcal{F}_t)_{t\in[0,T]}$*: If* $0 \leq s \leq t \leq T$*, we have* $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$.

We interpret $\mathcal{F}_t$ as the information known at time t. A probability space equipped with a filtration is called a *filtered probability space.*

We can divide stochastic processes into different classes, such as martingales, semimartingales, Brownian motion, Poisson processes and Lévy processes. Before defining the Lévy process, it is necessary to define some other stochastic processes. First we will give a definition of what it means that a process is càdlàg.

Definition 2.15 (Càdlàg). *A stochastic process* X *is said to be càdlàg if it is almost surely right-continuous with left-hand limits.*

Definition 2.16 (Martingale). *A càdlàg process* $(X_t)_{t\in[0,T]}$ *is a martingale if* X *is non-anticipating (adapted to* $\mathcal{F}_t$*),* $\mathrm{E}[|X_t|]$ *is finite for all* $t \in [0, T]$ *and*

$$\forall s > t,\ \mathrm{E}[X_s|\mathcal{F}_t] = X_t.$$

Definition 2.17 (Brownian motion). *An almost surely continuous process* $(X_t)_{t\in[0,T]}$ *is a Brownian motion if* $X_0 = 0$ *and* X_t *has independent increments, with* $X_t - X_s \sim \mathcal{N}(0, t-s)$*, where* $\mathcal{N}(\mu, \sigma^2)$ *denotes the normal distribution with expected value* μ *and variance* σ^2.

A Brownian motion can also be called a Wiener process.

Definition 2.18 (Lévy process). *A Lévy process is a continuous-time càdlàg stochastic process* $(L_t)_{t\geq 0}$ *with values in* $\mathbb{R}^d$*, such that*

1. $L_0 = 0$ *almost surely,*
2. *for any increasing sequence of times* $t_0, t_1, \ldots, t_n$*, the random variables* $L_{t_0}, L_{t_1} - L_{t_0}, \ldots, L_{t_n} - L_{t_{n-1}}$ *are independent (independent increments),*
3. *the law of* $L_{t+h} - L_t$ *does not depend on* t *(stationary increments), and*
4. $\forall \epsilon > 0, \forall t \geq 0, \lim_{h\to 0} \mathbb{P}(|L_{t+h} - L_t| \geq \epsilon) = 0$ *(stochastic continuity).*

Now we define the measure ν on $\mathbb{R}^d$ to be such that $\nu(A)$ is the expected number of jumps, per unit time, whose size belong to A. We also define a random measure N to be the number of actual jumps in a given time period of a given size.

Definition 2.19 (Lévy measure and jump measure). *Let* L_t *be a Lévy process on* $\mathbb{R}$*. Then the measure* ν *on* $\mathbb{R}$ *defined by*

$$\nu(A) = \mathrm{E}\big[\#\{t \in [0, 1] : \Delta L_t \neq 0, \Delta L_t \in A\}\big]$$

for all $A \in \mathcal{B}(\mathbb{R})$ *is called the Lévy measure of* L*. The random measure* N *on* $[0,\infty)\times\mathbb{R}$ *defined by*

$$N(B) = \#(t, L_t - L_{t^-}) \in B$$

for all $B \in \mathcal{B}\big([0,\infty)\times\mathbb{R}\big)$*, is called the jump measure of* L*.*

We will give a formula that shows how a Lévy process can be decomposed into different parts, but first we need to define what we mean by a Poisson random measure.

Definition 2.20 (Poisson random measure). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* $E \subseteq \mathbb{R}$ *and* μ *a given (positive) Radon measure on* $(E, \mathcal{E})$*. A Poisson random measure on* $(E, \mathcal{E})$ *with intensity measure* μ*, is an integer-valued random measure* $M : \Omega\times E \to \mathbb{N}$ *such that the following conditions are satisfied:*

1. *For (almost all)* $\omega \in \Omega$*,* $M(\omega,\cdot)$ *is an integer-valued Radon measure on* E*: For any bounded measurable* $A \subseteq E$*,* $M(A) < \infty$ *is an integer-valued random variable.*
2. *For each measurable set* $A \subseteq E$*,* $M(\cdot, A) = M(A)$ *is a Poisson random variable with parameter* $\mu(A)$*:*
$$\mathbb{P}\big(M(A) = k\big) = e^{\mu(A)}\frac{\big(\mu(A)\big)^k}{k!}$$
for all $k \in \mathbb{N}$*.*
3. *For disjoint measurable sets* $A_1, \ldots, A_n \subset E$*, the variables* $M(A_1), \ldots, M(A_n)$ *are independent.*

The following theorem gives a useful representation of Lévy processes.

Theorem 2.21 (Lévy-Itô Decomposition). *Let* L_t *be a Lévy process on* $\mathbb{R}$ *and* ν *its Lévy measure. Then* ν *is a Radon measure and verifies*

$$\int_{|x|\le 1} |x|^2\nu(dx) < \infty \qquad \text{and} \qquad \int_{|x|\ge 1} \nu(dx) < \infty.$$

The jump measure of L*, denoted by* N*, is a Poisson random measure on* $[0,\infty)\times\mathbb{R}$ *with intensity measure* $\nu(dx)dt$*. A Lévy process can be written as*

$$L_t = bt + \sigma B_t + \int_0^t \int_{|x|\ge 1} x\, N(ds,dx) + \lim_{\epsilon\to 0}\int_0^t \int_{\epsilon<|x|<1} x\, \tilde{N}(ds,dx), \tag{2.1}$$

where B_t *is a Brownian motion,* $b \in \mathbb{R}$*,* $\sigma > 0$ *and* $\tilde{N}(ds,dz) = N(ds,dx) - \nu(dx)ds$*. All terms of* (2.1) *are independent, and the convergence in the last term is almost sure and uniform in* t *on* $[0,T]$*.*

The last term on the right-hand side of (2.1) will be written as

$$\int_0^t \int_{0<|x|<1} x\, \tilde{N}(ds,dx).$$

We see that a Brownian motion is a Lévy processes.

Itô's formula is a useful tool when working with stochastic processes. It is the stochastic calculus counterpart of the chain rule in ordinary calculus.

Theorem 2.22 (Itô's formula). *Let $(X_t)_{t\geq 0}$ be a d-dimensional semimartingale. For any function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ that is continuously differentiable with respect to time, and twice continuously differentiable with respect to the other variables,*

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t f_s(s, X_s)\, ds + \sum_{i=1}^d \int_0^t f_i(s, X_s)\, dX_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t f_{ij}(s, X_{s^-})\, d[X, X]_s^c \\ &+ \sum_{\substack{0\leq s\leq t \\ \Delta X_s \neq 0}} \left(f(s, X_s) - f(s, X_{s^-}) - \sum_{i=1}^d \Delta X_s f_i(s, X_{s^-}) \right) \end{aligned}$$

for all $t \in [0,T]$, where $[X,X]^c$ denotes the continuous part of the quadratic variation process of X.

A Lévy process is a semimartingale, and therefore Itôs formula applies.

See [14] for more information about jump processes, and see [44] for more information about stochastic analysis in general.

We end the section with a theorem of probability, which will be used several times throughout the thesis. It follows directly from Theorem 34.4 of [11], which says that $\mathrm{E}[X] = \mathrm{E}\big[\mathrm{E}[X \mid \mathcal{F}_1]\big]$ for any σ-algebra $\mathcal{F}_1 \subset \mathcal{F}$. The theorem below can be proved by defining $\mathcal{F}_1$ as the σ-algebra generated by $\{\Omega_1, \Omega_2, \ldots\}$.

Theorem 2.23 (Partition Theorem of Conditional Expectation). *Let X be a random variable of finite expectation, and let $\{\Omega_n\}_{n\in\mathbb{N}}$ be a countable partition of Ω, i.e., $\Omega_n \cap \Omega_m = \emptyset$ for all $n \neq m$, and $\Omega = \cup_{n=1}^\infty \Omega_n$. Then*

$$\mathrm{E}[X] = \sum_{n=1}^\infty \mathrm{E}\big[X \mid \Omega_n\big] \cdot \mathbb{P}[\Omega_n].$$

2.3 Control theory and dynamic programming

Optimal control theory deals with the problem of finding a control law for a given system, such that a certain optimality criterion is achieved. A control problem includes a utility function that is a function of the state and control variables involved.

Control problems can be either deterministic or stochastic. In the first case the development of the state variables can be deterministically determined from the control; in the second case their development is a stochastic process. A deterministic control problem in continuous-time and finite time-horizon typically takes the form

$$V(t,x) = \sup_{C\in\mathscr{A}_{t,x}} \int_t^T U(s, X_s^C, C_s)\, ds + W(X_T^C, C_T),$$

while a stochastic control problem typically takes the form

$$V(t,x) = \sup_{C\in\mathscr{A}_{t,x}} \mathrm{E}\left[\int_t^T U(s, X_s^C, C_s)\,ds + W(X_T^C, C_T)\right]. \tag{2.2}$$

In these equations

- t represents time, and T is the terminal time.
- X_s^C is a vector that represents the state of the system at time s. The process X_s^C has initial value x, and develops according to the chosen control, i.e., $dX_s = f(X_s^C, C_s)$ for some function f.
- C_s is the value of the control at time s. The control C must belong to $\mathscr{A}_{t,x}$, the set of all admissible controls from the starting position (t, X).
- U and W are utility functions that give a measure of the satisfaction that is obtained in different states and for different controls.
- The value function V is a function describing the largest possible (expected) satisfaction that can be obtained from the starting state (t,x). The largest possible (expected) satisfaction is obtained by choosing an optimal control C^*.

The problem is to find the value function, in addition to determining an optimal control C^*. If possible, we wish to find an explicit formula for V and C^*, but in most cases we must find V and C^* by numerical methods.

HARA utility functions is a large class of utility functions, and CRRA utility functions is a subclass of HARA utility functions. A utility function $U \in C^2(\mathbb{R})$ exhibits HARA if its absolute risk aversion can be written on the form $1/(ac+b)$ for constants $a, b \in \mathbb{R}$, where the absolute risk aversion is defined by

$$-\frac{U''(c)}{U(c)}.$$

A utility function exhibits CRRA if its relative risk aversion, defined by

$$-\frac{cU''(c)}{U'(c)},$$

is constant. All CRRA functions can be written on the form $U(c) = Ac^\gamma/\gamma$ or $U(c) = A\ln c$ for some constant $A \in \mathbb{R}$. The utility function $U(c) = Ac^\gamma/\gamma$ corresponds to a relative risk aversion of $1-\gamma$, while $U(c) = A\ln c$ corresponds to a relative risk aversion of 1.

The Merton Problem is a famous example of a stochastic optimal control problem. An investor has initial wealth x, and the problem is to determine what fraction of the wealth that should be invested in a portfolio, and what fraction that should be invested in a risk-free asset. The uncertain asset follows a geometric Brownian motion with higher expected return than the risk-free asset, which has constant rent. The investor may transfer money between the risk-free asset and the stock at any time, and may also consume wealth. She obtains satisfaction from consumption and from wealth at the terminal time. This thesis concerns a generalization of the Merton problem.

Optimal control problems are satisfying the dynamic programming principle, which can be stated as follows for the stochastic control problem (2.2):

$$V(t,x) = \sup_{C\in\mathscr{A}_{t,x}} \mathrm{E}\left[\int_t^{t+\Delta t} U(s, X_s, C_s)\,ds + V\big(t+\Delta t, X_{t+\Delta t}^C\big)\right]$$

for all $\Delta t \in [0, T-t]$. The dynamic programming equation can be transformed to a so-called Hamilton-Jacobi-Bellman (HJB) equation equation by letting $\Delta t \to 0$:

$$F(t, x, V, DV, D^2V) = 0. \tag{2.3}$$

The HJB equation is typically a non-linear partial differential equation of first or second order, where the value function V is the unknown. If only continuous sample paths X_s are allowed, the equation often takes the form $V_t + \sup_c \bar{F}(t, x, D_xV, D_{xx}V, c) = 0$ for some function $\bar{F}$. If singular sample paths are allowed, the HJB equation is often a system of equations or inequalities instead of only a single equation. If the stochastic processes involved are jump processes, for example Lévy processes, the HJB equation will be non-local.

See [25] for more information about control theory.

2.4 Viscosity solutions of partial differential equations

Often the value function V is not sufficiently smooth to satisfy the HJB equation (2.3) in a classical sense, for example there may be points where V is not differentiable. See Chapter II.2 in [25] for examples of control problems with non-smooth value functions. One has introduced a generalized kind of solution to the HJB equation, called a *viscosity solution*, to overcome this problem.

A viscosity solution of (2.3) does not need to be differentiable everywhere, but it still satisfies (2.3) in an appropriate sense. See Definition 6.2 for a precise definition of viscosity solutions for the problem we will consider in this thesis. One can show that a smooth function satisfying the HJB equation is a viscosity solution, and that a smooth value function will satisfy the HJB equation in a classical sense.

There are many advantages with using viscosity theory, instead of only considering classical solutions of (2.3). First, it widens the class of problems we may consider, because we also can study problems where the value function is not differentiable. Second, it is often relatively easy to show existence and uniqueness of viscosity solutions of the HJB equation, and to show that the value function is a viscosity solution of (2.3). Third, we get precise formulations of general boundary conditions. Fourth, the theory often provides a great flexibility in passing to limits in various settings. Fifth, it is often easy to show continuous dependence of the value function V on the problem data, and it is sometimes also possible to show higher order regularity of V.

The concept of viscosity solutions is also applied in other fields than optimal control, for example in front evolution problems and differential games. See [25] for a good introduction to viscosity theory for controlled Markov (memoryless) processes, and see [15] for a good general introduction to viscosity solutions.

2.5 Penalty approximations

Penalty methods are a wide class of methods where one approximates a problem (P) by a family of problems $(P_\epsilon)_{\epsilon>0}$. Each problem (P) or (P_ϵ) consists of a partial differential equation (PDE), possibly in addition to some boundary values and additional constraints. We wish to find a family of penalty approximation $(P_\epsilon)_{\epsilon>0}$, such that the solution of (P_ϵ) converges to the solution of (P) when $\epsilon \to 0$. Penalty methods are useful when constructing numerical schemes for problems that are difficult to discretize. If (P) is difficult to solve numerically, while (P_ϵ) is easy to solve numerically, we solve the problem (P_ϵ) for a small value of ϵ, instead of solving (P) directly.

The form of the penalty approximation (P_ϵ) depends of course strongly on the original problem (P). We will give a few examples on how appropriate penalty approximations may be found. In [10] a penalty approximation is obtained by replacing a discontinuous function in the PDE by a continuous approximation. Another strategy for constructing a penalty method is presented in [43]. The authors first find a finite difference scheme for (P) on the form $\min\{A_c x = b_c : c \in C\}$, where A_c is a known matrix, b_c is a known vector, and C is a finite set of controls. Then they find a penalty approximation to the finite difference scheme on the form

$$(A_{c_0} x_\epsilon - b_{c_0}) - \epsilon \sum_{c \in C \setminus \{c_0\}} \max\{b_c - A_c x_\epsilon; 0\} = 0. \tag{2.4}$$

The solution of the penalty approximation is shown to converge to the solution of the original discrete problem.

In some cases (P) is difficult to solve because of additional constraints, and in this case one can modify the PDE of (P) by adding so-called penalty terms. One special case is when the PDE of (P) is on the form $\max\{F, G_1, G_2, \ldots, G_n\} = 0$ for some functions F, G_1, G_2,...,G_n of the unknown function and its derivatives. One possible penalty approximation is

$$F + f_1^\epsilon(G_1) + f_2^\epsilon(G_2) + \ldots + f_n^\epsilon(G_n) = 0, \tag{2.5}$$

where $\lim_{\epsilon \to 0} f_i^\epsilon(G_i) = 0$ if $G_i < 0$, $f_i^\epsilon(G_i) \geq 0$ if $G_i = 0$, and $f_i^\epsilon(G_i)$ blows up as $\epsilon \to 0$ if $G_i > 0$. See [38], [17] and [16] for problems where a penalty approximation on the form (2.5) is used. A similar idea is used in [32], though the penalty approximation cannot be written explicitly on the form (2.5) in this case. It is easy to see intuitively that the solution of (2.4) and (2.5) may converge to the solution of the original problem as $\epsilon \to 0$. However, giving a careful proof that the solutions of (2.4) and (2.5) do not blow up as $\epsilon \to 0$, and that the limiting function satisfies the original problem, is more challenging. In Chapter 10 it is shown that a penalty approximation on the form (2.5) gives a uniformly convergent sequence of solutions for the problem considered in this thesis.

2.6 Analysis

In this section some known theorems of analysis that will be used in later chapters, are stated. Basic definitions of for example uniform continuity are assumed known to the reader. The reader is referred to [20], [3], [22], [23] and [34] for proofs.

Definition 2.24 (Equicontinuity). *Let $\{f_k\}_{k\in\mathbb{N}}$ be a sequence of functions defined on $A \subset \mathbb{R}^n$. They are said to be equicontinuous if, for each $\epsilon > 0$ and $x \in A$, there is a $\delta > 0$ such that $|y - x| < \delta$ and $y \in A$ implies $|f_k(x) - f_k(y)| < \epsilon$ for all $k \in \mathbb{N}$.*

The Arzelà-Ascoli theorem can be used to prove that a sequence of functions converge uniformly.

Theorem 2.25 (Arzelà-Ascoli Theorem). *If a sequence $\{f_n\}_{n\in\mathbb{N}}$ of continuous functions defined on a compact set is bounded and equicontinuous, then $\{f_n\}_{n\in\mathbb{N}}$ has a subsequence converging uniformly to a continuous function f.*[1]

The Heine-Cantor theorem gives us a condition that guarantees that a function f is uniformly continuous.

Theorem 2.26 (Heine-Cantor Theorem). *Let $f : A \to B$ be a continuous function defined on a compact set $A \subset \mathbb{R}^n$ taking values in the set $B \subset \mathbb{R}^m$. Then f is uniformly continuous.*

A function is uniformly continuous if and only if it admits a *modulus of continuity.*

Theorem 2.27. *The function $v : A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$, is uniformly continuous if and only if there exists a function $\omega : \mathbb{R}^n \to \mathbb{R}$ such that $\omega(0,0,0) = 0$, ω is continuous at $(0,0,0)$ and*

$$|v(t,x,y) - v(t',x',y')| \leq \omega(|t-t'|, |x-x'|, |y-y'|)$$

for all $(t,x,y), (t',x',y') \in A$.

The function ω is called the modulus of continuity of v.

Grönwall's inequality can be used to show bounds on a function defined by an integral of itself.

Theorem 2.28 (Grönwall's inequality). *Let f be a non-negative integrable function on $[0,T]$ that satisfies the integral inequality*

$$f(t) \leq C_1 \int_0^t f(s)\,ds + C_2$$

for almost all $t \in [0,T)$ and constants $C_1, C_2 \geq 0$. Then

$$0 \leq f(t) \leq C_2 e^{C_1 t}$$

for a.e. $t \in [0,T]$.

Theorem 2.29 (Banach fixed point theorem). *Let $f : A \to A$, $A \subseteq \mathbb{R}^n$, and suppose there is a constant $\epsilon > 0$ such that*

$$|f(x) - f(y)| \leq (1-\epsilon)|x-y|$$

for all $x, y \in A$. Then there exist a unique point $a \in A$ such that $f(a) = a$. For fixed $x \in X$, the sequence $\{f^n(x)\}_{n\in\mathbb{N}}$ converges to a, where the functions $f^n : A \to A$ are defined recursively by $f^0(x) = x$ and $f^{n+1}(x) = f\big(f^n(x)\big)$ for $n \in \mathbb{N}$.

[1] The most common formulation of the Arzelà-Ascoli Theorem says that the functions $\{f_n\}_{n\in\mathbb{N}}$ should be *uniformly* equicontinuous, but as shown in [20], the formulation given here is equivalent.

Part I
A singular stochastic portfolio optimization problem

In this part a singular stochastic portfolio optimization problem is presented, and various properties of the problem will be discussed and proved in a viscosity solution setting. Many of the results are adaptions of results proved in [8] and [9] for the case of an infinite time-horizon.

In Chapter 3 the problem and the conditions on the functions and constants involved are stated, and in Chapter 4 some properties of the value function are proved. In Chapter 5 the HJB equation of the problem is derived by a heuristic argument, and it is shown that the dimension of the problem can be reduced for one specific class of utility functions. In Chapter 6 it is proved that the value function is the unique constrained viscosity solution of the HJB equation, and in Chapter 7 some explicit solution formulas are found.

Chapter 3
A singular stochastic portfolio optimization problem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t\in[0,T]}$ a given filtration. Consider a financial market that consists of a bond with constant interest rate $\hat{r} > 0$, and a stock that has price dynamics $S = (S_t)_{t\in[0,T]}$ given by

$$S_t = S_0 e^{L_t},$$

where $S_0 > 0$ and $L = (L_t)_{t\in[0,T]}$ is a Lévy process. The Lévy-Itô decomposition of L (see Theorem 2.21) is

$$L_t = \mu t + \sigma B_t + \int_0^t \int_{\{|z|<1\}} z\,\tilde{N}(ds, dz) + \int_0^t \int_{\{|z|\geq 1\}} z\,N(ds, dz),$$

where $\mu \in \mathbb{R}$, $\sigma \geq 0$, $B = (B_t)_{t\in[0,T]}$ is a Brownian motion and $N(dt, dz)$ is the jump measure of L with a deterministic compensator of the form $\nu(dz) \times dt$. The measure ν is the Lévy measure of L, see Definition 2.19, and ν satisfies

$$\nu(\{0\}) = 0, \int_{\mathbb{R}\setminus\{0\}} (|z|^2 \wedge 1)\,\nu(dz) < \infty \text{ and } \int_{|z|\geq 1} |e^z - 1|\,\nu(dz) < \infty. \tag{3.1}$$

The standard model of stock prices is the geometric Brownian motion, but a Lévy process is used instead of a Brownian motion here, as this is a more realistic model.

Using Itô's Formula, we get

$$dS_t = \hat{\mu} S_t\,dt + \sigma S_t\,dB_t + S_{t^-} \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\,\tilde{N}(dt, dz),$$

where

$$\hat{\mu} = \mu + \frac{1}{2}\sigma^2 + \int_{\mathbb{R}\setminus\{0\}} (e^z - 1 - z\mathbf{1}_{|z|<1})\,\nu(dz).$$

We assume $\hat{\mu} > \hat{r}$, i.e., the expected return from the stock is higher than the return of the bond. The advantage of investing in the stock instead of the bond, is that the expected return is higher. The disadvantage is that the risk is higher.

Consider an investor who wants to put her money in the stock and the bond, with the goal of maximizing her utility over some time interval $[t, T]$, $t \in [0, T]$. Let $\pi_s \in [0, 1]$ be the fraction of her wealth invested in the stock at time $s \in [t, T]$, and assume that there are no transaction costs in the market. Let C_s denote the cumulative consumption from time t up to time s, and let $X_s^{\pi,C}$ denote the wealth process. We have

$$\begin{aligned} X_s^{\pi,C} =& x - C_s + \int_t^s (\hat{r} + (\hat{\mu} - \hat{r})\pi_{s'}) X_{s'}^{\pi,C}\, ds' + \int_t^s \sigma \pi_{s'} X_{s'}^{\pi,C}\, dB_{s'} \\ &+ \int_t^s \pi_{s'-} X_{s'-}^{\pi,C} \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\, \tilde{N}(ds, dz), \end{aligned} \tag{3.2}$$

where x is the wealth at the starting time t, and π and C are the controls. To incorporate the idea of durability and local substitution, $Y_s^{\pi,C}$ is introduced. The process $Y_s^{\pi,C}$ represents the average past consumption, and is defined by

$$Y_s^{\pi,C} = ye^{-\beta(s-t)} + \beta e^{-\beta s} \int_t^s e^{\beta s'}\, dC_{s'}, \tag{3.3}$$

where y is the average past consumption at time t, and $\beta > 0$ is a weighting factor. The terms durability and local substitution will be explained below. The integrals in the expressions for $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ are interpreted pathwise in a Lebesgue-Stieltjes sense, see [30] for more information.

The goal of the investor is to find an allocation process π^* and a consumption plan C^* that optimize expected terminal wealth and expected discounted utility over the investment horizon $[t, T]$. Define the value function $V : \overline{\mathscr{D}_T} \to [0, \infty)$ by

$$V(t, x, y) = \sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right], \tag{3.4}$$

where $\mathscr{D}_T = \mathscr{D} \times [0, T)$, $\mathscr{D} = \left\{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\right\}$, $\delta > 0$ is the discount factor, $\mathscr{A}_{t,x,y}$ is a set of admissible controls, U and W are given utility functions, and E denotes the expectation. A value function on a form similar to this is common in the literature, see for example [31], where a more general function $U\left(s, Y_s^{\pi,C}\right)$ is used instead of $e^{-\delta s}U\left(Y_s^{\pi,C}\right)$, and [17], where U is a function of *present* consumption and $W = W\left(X_T^{\pi,C}\right)$ is on a form similar to U. The purpose of the discount factor δ is to signalize that we value early consumption more than late consumption.

To simplify notation, define $X := (x, y)$, $V(t, X) := V(t, x, y)$, $\mathscr{A}_{t,X} := \mathscr{A}_{t,x,y}$, etc. To distinguish $X = (x, y)$ from the stochastic process $X_s^{\pi,C}$, $X_s^{\pi,C}$ will always be marked with the controls π, C, unless something else is explicitly stated. In some cases it is most convenient to work with X on the full form (x, y), in other cases it is easier to use the short-hand notation X.

For any $(t, x, y) \in \overline{\mathscr{D}_T}$, $(\pi, C) \in \mathscr{A}_{t,x,y}$ if the following conditions hold:

(c_i) C_s is an adapted process on $[t, T]$ that is non-decreasing and right-continuous with left-hand limits (càdlàg). It has initial value $C_{t^-} = 0$ (to allow an initial jump when $C_t > 0$) and satisfies $\mathrm{E}[C_s] < \infty$ for all $s \in [t, T]$.
(c_{ii}) π_s is an adapted càdlàg process with values in $[0, 1]$.
(c_{iii}) $X_s^{\pi,C} \geq 0$, $Y_s^{\pi,C} \geq 0$ almost everywhere for all $s \in [t, T]$.

The restriction $\pi_s \leq 1$ for all $s \in [t, T]$ implies that the investor cannot loan money with rent $\hat{r}$ to invest in the risky asset. The restriction $X_s^{\pi,C} \geq 0$ means that the investor cannot spend more money than she has. The following assumptions are made on the utility functions $U : [0, \infty) \to [0, \infty)$ and $W : \overline{\mathscr{D}} \to [0, \infty)$:

(u_i) U and W are continuous, non-decreasing and concave.

(u_{ii}) There exist constants $K_U > 0$, $K_W > 0$ and $\gamma \in (0,1)$ such that $\delta > k(\gamma)$,

$$U(y) \leq K_U(1+y)^\gamma,$$

and

$$W(x,y) \leq K_W(1+x+y)^\gamma,$$

where

$$k(\gamma) := \max_{\pi\in[0,1]} \Bigg[\gamma\big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big) + \frac{1}{2}\gamma(\gamma-1)(\sigma\pi)^2 \\ + \int_{\mathbb{R}\setminus\{0\}} \big(1 + \pi(e^z - 1)\big)^\gamma - 1 - \gamma\pi(e^z-1)\,\nu(dz) \Bigg].$$

That U and W are concave, implies that the investor prioritizes safe investments over risky investments. If $X^{\pi,C}$ and $Y^{\pi,C}$ decrease, the satisfaction of the investor decreases, but if $X^{\pi,C}$ and $Y^{\pi,C}$ are increased by the same amount, the satisfaction is not increased by a corresponding amount. If U and W had been linear, the investor had not given priority to safe investments over risky investments; her only goal had been to maximize the expected values of $X^{\pi,C}$ and $Y^{\pi,C}$.

A lemma showing that k is well-defined, will be proved below, but first we will define $\gamma^* \in (0,1)$ and the sets $C_l(\overline{\mathscr{D}_T})$ and $C'_l(\overline{\mathscr{D}_T})$.

Definition 3.1. *Define*

$$\gamma^* = \sup\{\gamma \in (0,1] : \delta > k(\gamma)\}.$$

Note that $\delta = k(\gamma^*)$, since k is continuous.

Definition 3.2. *For all $l \geq 0$, define*

$$C_l(\overline{\mathscr{D}_T}) = \left\{ \phi \in C(\overline{\mathscr{D}_T}) : \sup_{\overline{\mathscr{D}_T}} \frac{|\phi(t,x,y)|}{(1+x+y)^l} < \infty \right\}.$$

Note that $C_{l_1}(\overline{\mathscr{D}_T}) \subset C_{l_2}(\overline{\mathscr{D}_T})$ if $l_1 < l_2$.

Definition 3.3. *For all $l \geq 0$, define*

$$C'_l(\overline{\mathscr{D}_T}) = \big\{\phi \in C(\overline{\mathscr{D}_T}) \,:\, \phi \in C_{l'}(\overline{\mathscr{D}_T}) \text{ for some } l' < l\big\}.$$

Note that $C'_l(\overline{\mathscr{D}_T}) \subset C_l(\overline{\mathscr{D}_T})$, and that $C_{l'}(\overline{\mathscr{D}_T}) \subset C'_l(\overline{\mathscr{D}_T})$ for all $l' < l$. We will prove uniqueness of viscosity solutions in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$ in a later chapter. The following lemma shows that k is well-defined (putting $\hat{V}(t,x,y) = x^\gamma$ and $x = 1$), and will also be used in later chapters.

Lemma 3.4. *Let $\widehat{V} \in C_1(\overline{\mathscr{D}_T}) \cap C^{1,2,1}(\overline{\mathscr{D}_T})$. Then*

$$\int_{\mathbb{R}\setminus\{0\}} \widehat{V}\big(t, x + \pi x(e^z-1), y\big) - \widehat{V}(t,x,y) - \pi x(e^z-1)\widehat{V}_x(t,x,y)\,\nu(dz) < \infty$$

for all $(t,x,y) \in \overline{\mathscr{D}_T}$.

Proof. Fix $(t,x,y) \in \overline{\mathscr{D}_T}$. Since $\widehat{V} \in C_1(\overline{\mathscr{D}_T})$, there is a constant M, depending on (t,x,y), such that

$$\left|\widehat{V}(t,x+\pi x(e^z-1),y) - \widehat{V}(t,x,y) - \pi x(e^z-1)\widehat{V}_x(t,x,y)\right| \le M(1+|e^z-1|)$$

for all $|z|>1$. By (3.1) and the dominated convergence theorem (Theorem 2.8), we see that the integral is defined for $|z|>1$.

For $|z| \le 1$, we make a Taylor expansion of $\widehat{V}$ around (t,x,y):

$$\begin{aligned}\widehat{V}\big(t,x+\pi x(e^z-1),y\big) - \widehat{V}(t,x,y) - \pi x(e^z-1)\widehat{V}_x(t,x,y)&\\ = \frac{1}{2}\widehat{V}_{xx}(t,a,y)\big(\pi x(e^z-1)\big)^2&\\ = O(z^2),&\end{aligned}$$

where a is between x and $x+x\pi(e^z-1)$. Again we see by (3.1) that the integral is finite. □

Our problem is said to exhibit durability and local substitution, because U is a function of average past consumption, instead of present consumption rate, see [31]. Durability means that the satisfaction the agent gets from consumption is lasting, and C_t can be interpreted as the total purchase of a durable good until time t, where the purchase of the durable good is irreversible. Local substitution means that consumption at nearby dates are almost perfect substitutes. The standard Merton problem with consumption does not exhibit this property. It is not explained why this is the case in [31], but we will give a short explanation of why our problem exhibit local substitution, while the Merton problem does not:

If an agent consumes at rate c_1 in the time interval $[t,t+\Delta t]$, and at rate c_2 in the time interval $[t+\Delta t, t+2\Delta t]$, her total satisfaction from consumption in $[t,t+2\Delta t]$ is approximately equal to

$$2\Delta t e^{-\delta t}\big(U(c_1)+U(c_2)\big),$$

if she only obtains satisfaction from present consumption. If she consumes at rate $(c_1+c_2)/2$ in the interval $[t,t+2\Delta t]$, on the other hand, her total satisfaction is

$$2\Delta t e^{-\delta t} U\left(\frac{c_1+c_2}{2}\right).$$

We see that her satisfaction is not the same in the two cases, since U is not a linear function. Since U is a concave function, the agent will obtain largest satisfaction in the last case, i.e., the case where the consumption rate is constant. However, her total consumption is equal to $(c_1+c_2)\Delta t$ in both cases. If U is a function of average past consumption, $Y^{\pi,C}$ will be approximately constant in the interval $[t,t+2\Delta t]$ for sufficiently small Δt, and $Y_T^{\pi,C}$ will be approximately identical for the two cases, see (3.3). Therefore the agent's satisfaction will be approximately equal in the two cases if U is a function of average past consumption.

We assume the dynamic programming principle is valid. The dynamic programming principle can be stated formally as in the following theorem.

Theorem 3.5 (The Dynamic Programming Principle). *For all* $(t,x,y) \in \overline{\mathscr{D}_T}$ *and* $\Delta t \in [0, T-t]$, *we have*

$$V(t,x,y) = \sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^{t+\Delta t} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds + V\left(t+\Delta t, X_{t+\Delta t}^{\pi,C}, Y_{t+\Delta t}^{\pi,C}\right)\right].$$

We can easily see that the dynamic programming principle is valid from a probabilistic point of view, but a proof of its validity will not be given here. A proof of the dynamic programming principle for another stochastic optimization problem with jump processes is given in [33]. See also [25] and [24] for proofs of the dynamic programming principle for various control problems. The existence of optimal controls will also be assumed. This is proved for a slightly different singular stochastic control problem in [7]. The existence of an optimal control for our problem, is stated in the following theorem.

Theorem 3.6. *For all* $(t,x,y) \in \overline{\mathscr{D}_T}$ *there exists* $(\pi^*, C^*) \in \mathscr{A}_{t,x,y}$ *such that*

$$V(t,x,y) = \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi^*,C^*}\right)\, ds + W\left(X_T^{\pi^*,C^*}, Y_T^{\pi^*,C^*}\right)\right].$$

3.1 Notation

In addition to standard notation and the notation introduced above, the following definitions will be used throughout the thesis:

$C^{k,l,m}(\mathscr{O}_T)$, $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$, denotes the set of all functions $f : \mathscr{O}_T \to \mathbb{R}$ that are k times continuously differentiable in t, l times continuously differentiable in x, and m times continuously differentiable in y.

For all $A \subseteq \mathbb{R}^n$, $USC(A)$ $\big(LSC(A)\big)$ denotes the set of all upper semi-continuous (lower semi-continuous) functions defined on A.

If $f, g : A \to \mathbb{R}$ for some $A \subset \mathbb{R}^n$, the function $h = \max\{f; g\}$ is defined as the pointwise supremum of the two functions f and g.

Let $A \subseteq \mathbb{R}^n$, and let B be a boolean expression that is either satisfied or not satisfied by each element of A. Define $\mathbf{1}_B : A \to \{0,1\}$ to be such that $(\mathbf{1}_B)(a) = 1$ if a satisfies B, and $(\mathbf{1}_B)(a) = 0$ if a does not satisfy B.

$\mathscr{N}\big((t,X), r\big)$, $(t,X) \in \overline{\mathscr{D}_T}$, $r > 0$ denotes the ball with centre (t,X) and radius r, i.e.,

$$\mathscr{N}\big((t,X), r\big) := \big\{(s,Y) \in \mathbb{R}^3 : |(s,Y) - (t,Y)| < r\big\}.$$

We let $\overline{\mathscr{N}}\big((t,X), r\big)$ denote the closure of $\mathscr{N}$.

$\mathbb{S}^n$ denotes the set of all $n \times n$ symmetric matrices, where $n \in \mathbb{N}$.

If $A \in \mathbb{S}^n$, we let $|A|$ denote the spectral radius norm of A.

If $A, B \in \mathbb{S}^n$, we let $B \leq A$ mean that $A - B$ is positive semidefinite.

I^n denotes the $n \times n$ identity matrix. If the dimension of the matrix is obvious from the context, the superscript n may be skipped.

If $f \in C^1(A)$ for some $A \subseteq \mathbb{R}^n$, Df denotes the vector of all partial derivatives of f:

$$Df = \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{pmatrix},$$

where f_i is the derivative of f with respect to the ith argument. If X is a vector of input variables, $D_X f$ denotes the part of Df associated with X.

If $f \in C^2(A)$ for some $A \subseteq \mathbb{R}^n$, $D^2 f$ denotes the vector of all second partial derivatives of f:

$$D^2 f = \begin{pmatrix} f_{1,1} & f_{1,2} & \dots & f_{1,n} \\ f_{2,1} & f_{2,2} & \dots & f_{2,n} \\ \dots & \dots & & \dots \\ f_{n,1} & f_{n,2} & \dots & f_{n,n} \end{pmatrix},$$

where $f_{i,j}$ is the derivative of f with respect to the ith and the jth argument. If X is a vector of input variables, $D^2_X f$ denotes the part of $D^2 f$ associated with X. If X is a single variable, we may write $\partial^2_X f$ instead of $D^2_X f$. If X is a single variable, we may also use the notation $\partial^n_X f$ for higher-order derivatives of f with respect to X. This notation is easier than $f_X, f_{XX}, f_{XXX}, f_{XXXX}$ etc when the number of derivatives is large.

For each vector $v = (v_1\, v_2\, \dots\, v_n)^t \in \mathbb{R}^n$, $|v|$ denotes the Euclidean norm of v, i.e.,

$$|v| = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}.$$

For each function $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}^n$, $|f|_{L^\infty}$ denotes the supremum norm of f, i.e.,

$$|f|_{L^\infty} = \sup_{x \in A} |f(x)|.$$

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\Omega' \subseteq \Omega$, we let $\mathbb{P}[\Omega']$ denote the probability of Ω'. If $f : \Omega \to \mathbb{R}$ and $\Omega' \subset \Omega$, we let

$$E\Big[f(\omega) \,|\, \Omega'\Big]$$

denote the expected value of f given $\omega \in \Omega'$.

Chapter 4
Properties of the value function

In this chapter some properties of the value function are stated and proved. First it is proved that V is well-defined. Then we state the boundary conditions and the terminal value condition, and show that the terminal value condition may be rewritten on a form that will be more convenient in proofs in later chapters. Then we show some monotonicity and growth properties of V, and at the end of the chapter some regularity results are proved. Many of the proofs are following the same ideas as similar proofs in [9], where the authors consider an infinite-horizon version of our optimization problem with a pure-jump Lévy process. The HJB equation in the infinite-horizon case does not depend on t, so the results involving the time variable t are new.

Lemma 4.1. *The value function V is well-defined, i.e., the right-hand side of* (3.4) *exists for all $(t,x,y) \in \overline{\mathscr{D}_T}$.*

Proof. Since all sets of real numbers have a supremum (finite or ∞), it is sufficient to show that the set $\mathscr{A}_{t,x,y}$ is non-empty for all $(t,x,y) \in \overline{\mathscr{D}_T}$. Define $\pi \equiv \pi^*$ and $C \equiv 0$ for some $\pi^* \in [0,1]$. Then $(\pi, C) \in \mathscr{A}_{t,x,y}$ for all $(t,x,y) \in \overline{\mathscr{D}_T}$, because C is non-decreasing and $X_s^{\pi,C} \geq 0$ for all $s \in [t,T]$. □

The lemma above does not say that V takes a *finite* value for all $(t,x,y) \in \overline{\mathscr{D}_T}$. However, we can use Lemma 4.6 to prove this later.

4.1 Boundary conditions and terminal value condition

At the boundary of $\mathscr{D}$ we have state constraints. These say that $X_s^{\pi,C}$ and $Y_s^{\pi,X}$ must stay non-negative for all $s \in [t,T]$. We will see in Chapter 6 that this can be expressed mathematically as V satisfying a specific viscosity subsolution inequality at the boundary.

For $x = 0$ we can find an explicit expression for V. If $x = 0$, we have $C_s = 0$ and $X_s^{\pi,C} = 0$ for all $s \geq t$, because $X^{\pi,C}$ must stay non-negative. From (3.3) we see that $Y_s = ye^{-\beta(s-t)}$, and this implies that

$$V(t,0,y) = \int_t^T e^{-\delta s} U\left(ye^{-\beta(s-t)}\right) ds + W\left(0, ye^{-\beta(T-t)}\right). \tag{4.1}$$

For $y = 0$ it is reasonable to assume that V satisfies $\beta v_y - v_x = 0$ in a viscosity sense. This will be explained in Chapter 11.

If $t = T$, (3.2) and (3.3) give

$$X_T^{\pi,C} = x - C_T$$

and

$$Y_T^{\pi,C} = y + \beta e^{-\beta T} \int_{T^-}^{T} e^{\beta s}\, dC_s = y + \beta C_T.$$

The control C_s is only defined for $s = T$ in this case. The control C is admissible iff $0 \leq C_T \leq x$, where $C_T > 0$ corresponds to a consumption gulp. We get

$$V(T, x, y) = \max_{c \in [0,x]} W(x - c, y + \beta c). \tag{4.2}$$

We see from this expression that it might be possible to reformulate the problem with a more convenient definition of W. We call this function $\widehat{W}$, and it is defined as follows:

Definition 4.2. *Define* $\widehat{W} : \overline{\mathscr{D}} \to [0, \infty)$ *by*

$$\widehat{W}(x, y) := \max_{c \in [0,1]} W(x - c, y + \beta c)$$

for all $(x, y) \in \overline{\mathscr{D}}$.

We see from (4.2) that

$$V(t, x, y) = \sup_{(\pi,C) \in \mathscr{A}_{t,x,y}} \left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds + \widehat{W}\left(X_T^{\pi,C}, Y_T^{\pi,C}\right) \right],$$

because the optimal control always will do a terminal consumption gulp if this increases the value of W. We also see that $\widehat{W}$ satisfies the same growth conditions as W:

Lemma 4.3. *The function* $\widehat{W}$ *satisfies* $(u_i) - (u_{ii})$.

Proof. We see immediately that $\widehat{W}$ is continuous, since W is continuous. By using that W is non-decreasing, we see that $\widehat{W}$ also is non-decreasing, because, for any $\Delta x, \Delta y > 0$,

$$\begin{aligned}
\widehat{W}(x + \Delta x, y) &= \max_{c \in [0, x + \Delta x]} W(x + \Delta x - c, y + \beta c) \\
&\geq \max_{c \in [0,x]} W(x + \Delta x - c, y + \beta c) \\
&\geq \max_{c \in [0,x]} W(x - c, y + \beta c) \\
&= \widehat{W}(x, y)
\end{aligned}$$

and

$$\begin{aligned}
\widehat{W}(x, y + \Delta y) &= \max_{c \in [0,x]} W(x - c, y + \Delta y + \beta c) \\
&\geq \max_{c \in [0,x]} W(x - c, y + \beta c) \\
&= \widehat{W}(x, y).
\end{aligned}$$

Now we want to prove that $\widehat{W}$ is concave. Fix any $(x_1, y_1), (x_2, y_2) \in \overline{\mathscr{D}}$, and let $c_1 \in [0, x_1], c_2 \in [0, x_2]$ satisfy

$$\widehat{W}(x_i, y_i) = W(x_i - c_i, x_i + \beta c_i)$$

for $i = 1, 2$. The existence of c_1 and c_2 follows from the compactness of $[0, x_1]$ and $[0, x_2]$. By using the concavity of W, we have, for any $\lambda \in [0, 1]$,

$$\begin{aligned}
&\lambda\widehat{W}(x_1, y_1) + (1-\lambda)\widehat{W}(x_2, y_2) \\
&\quad = \lambda W(x_1 - c_1, x_1 + \beta c_1) + (1-\lambda)W(x_2 - c_2, x_2 + \beta c_2) \\
&\quad \leq W\big(\lambda x_1 + (1-\lambda)x_2 - \lambda c_1 - (1-\lambda)c_2, \lambda y_1 + (1-\lambda)y_2 + \beta(\lambda c_1 + (1-\lambda)c_2)\big) \\
&\quad \leq \max_{c \in [0, \lambda x_1 + (1-\lambda)x_2]} W(\lambda x_1 + (1-\lambda)x_2 - c, \lambda y_1 + (1-\lambda)y_2 + \beta c) \\
&\quad = \widehat{W}(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2).
\end{aligned}$$

We have now proved that $\widehat{W}$ satisfies (u_i). By using that W is non-decreasing and satisfies (u_{ii}), we see that $\widehat{W}$ satisfies a similar estimate as the one given in (u_{ii}):

$$\widehat{W}(x, y) \leq W(x, y + \beta x) \leq K_W\big(1 + (1+\beta)x + y\big)^\gamma < K_W(1+\beta)^\gamma(1 + x + y)^\gamma.$$

□

It is possible to reformulate the problem with the function $\widehat{W}$ instead of the function W. Some of the proofs in the thesis become easier with the new formulation of W. In Part II, where we define a non-singular version of the problem considered here, the value function V_ϵ will not be equal with the terminal utility function W and the terminal utility function $\widehat{W}$. If the terminal utility function is $\widehat{W}$, V_ϵ will converge to V on the whole domain $\overline{\mathscr{D}_T}$. If the terminal utility function is W, however, V_ϵ will only converge to V on $[0, T) \times \overline{\mathscr{D}}$.

4.2 Growth and Monotonicity Properties

In this section we will prove some growth and monotonicity properties of V. We start by considering the monotonicity of V in x and y. Since U and W are non-decreasing in x and y, it is reasonable that also V is non-decreasing in x and y.

Lemma 4.4. *The value function V is non-decreasing in x and y.*

Proof. Let $t \in [0, T]$, $(x, y), (x', y') \in \overline{\mathscr{D}}$ for $x \leq x'$ and $y \leq y'$, and $(\pi, C) \in \mathscr{A}_{t,x,y}$. $(\pi, C) \in \mathscr{A}_{t,x',y'}$, because $x' \geq x$ implies $X_s'^{\pi,C} \geq X_s^{\pi,C} \geq 0$, where $X'^{\pi,C}$ and $Y'^{\pi,C}$ are the processes with initial values x' and y', respectively, and $X^{\pi,C}$ and $Y^{\pi,C}$ are the processes with initial values x and y, respectively. It follows that $\mathscr{A}_{x,y,t} \subseteq \mathscr{A}_{t,x',y'}$. We also see that $Y_s^{\pi,C} \leq Y_s'^{\pi,C}$, and using the monotonicity of U and W (see (u_i)), we get

$$\begin{aligned}
V(t, x, y) &= \sup_{(\pi,C) \in \mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right] \\
&\leq \sup_{(\pi,C) \in \mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s'^{\pi,C}\right) ds + W\left(X_T'^{\pi,C}, Y_T'^{\pi,C}\right)\right] \\
&\leq \sup_{(\pi',C') \in \mathscr{A}_{t,x',y'}} \mathrm{E}\left[\int_t^T U\left(Y_s'^{\pi',C'}\right) ds + W\left(X_T'^{\pi',C'}, Y_T'^{\pi',C'}\right)\right] \\
&= V(t, x', y').
\end{aligned}$$

□

V is not generally increasing or decreasing in t. However, as the following examples show, we may manage to prove monotonicity of V in t for some special cases:

1. If $x = 0$ and
$$\int_{t_1}^{T} e^{-\delta s} U\left(ye^{-\delta(s-t_1)}\right) ds - \int_{t_2}^{T} e^{-\delta s} U\left(ye^{-\delta(s-t_2)}\right) ds$$
$$< W\left(0, ye^{-\beta(T-t_2)}\right) - W\left(0, ye^{-\beta(T-t_1)}\right)$$
for all $y > 0$ and $t_1, t_2 \in [0, T]$ with $t_1 < t_2$, V is increasing in t. If $x = 0$, we must have $C_s \equiv 0$ and $X_s^{\pi,C} \equiv 0$ for all $s \in [t, T]$, so $Y_s^{\pi,C} = ye^{-\beta(s-t)}$, and
$$\begin{aligned} V(t_1, 0, y) =& \int_{t_1}^{T} e^{-\delta s} U\left(ye^{-\delta(s-t_1)}\right) ds - W\left(0, ye^{-\beta(T-t_1)}\right) \\ <& \int_{t_2}^{T} e^{-\delta s} U\left(ye^{-\delta(s-t_2)}\right) ds + W\left(0, ye^{-\beta(T-t_2)}\right) \\ =& V(t_2, 0, y). \end{aligned}$$
One special case of this example is $U \equiv 0$, i.e., the case where the agent only is concerned about the terminal values of $X^{\pi,C}$ and $Y^{\pi,C}$.
2. If
$$W(x, y) \leq W(xe^{rt}, ye^{-\beta t})$$
for all $(t, x, y) \in \overline{\mathscr{D}_T}$, V is non-increasing in t. Let $t_1, t_2 \in [0, T]$ satisfy $t_1 \leq t_2$, and let $(\pi^*, C^*) \in \mathscr{A}_{t_2,x,y}$ be an optimal control. Define $(\pi, C) \in \mathscr{A}_{t_1,x,y}$ by
$$\pi_s = \begin{cases} \pi^*_{s-t_1+t_2} & \text{for } s \leq T - t_2 + t_1, \\ 0 & \text{for } s > T - t_2 + t_1, \end{cases}$$
and
$$C_s = \begin{cases} C^*_{s-t_1+t_2} & \text{for } s \leq T - t_2 + t_1, \\ C^*_T & \text{for } s > T - t_2 + t_1. \end{cases}$$
We get
$$\begin{aligned} V(t_1, x, y) \geq& \mathrm{E}\left[\int_{t_1}^{T} e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right] \\ \geq& \mathrm{E}\left[\int_{t_1}^{T+t_1-t_2} e^{-\delta s} U\left(Y_{T+t_1-t_2}^{\pi,C}\right) ds + W\left(X_{T+t_1-t_2}^{\pi,C}, Y_{T+t_1-t_2}^{\pi,C}\right)\right] \\ &+ \mathrm{E}\left[W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right) - W\left(X_{T+t_1-t_2}^{\pi,C}, Y_{T+t_1-t_2}^{\pi,C}\right)\right] \\ =& V(t_2, x, y) + \mathrm{E}\Big[W\left(X_{T+t_1-t_2}^{\pi,C} e^{\hat{r}(t_2-t_1)}, Y_{T+t_1-t_2}^{\pi,C} e^{-\beta(t_2-t_1)}\right) \\ &- W\left(X_{T+t_1-t_2}^{\pi,C}, Y_{T+t_1-t_2}^{\pi,C}\right)\Big] \\ \geq& V(t_2, x, y), \end{aligned}$$
for all $(t, x, y) \in \overline{\mathscr{D}_T}$. One special case of this example is the case $W \equiv 0$.

3. If $\sigma = 0$, $\nu\big((-\infty,0)\big) = 0$ and $U \equiv 0$, V is non-decreasing in t for sufficiently small values of $\hat{\mu}$ and $\hat{r}$. Since $\sigma = 0$ and $\nu\big((-\infty,0)\big) = 0$, $X^{\pi,C}$ is increasing in the absence of consumption, and since $U \equiv 0$, the agent is only concerned about the terminal values of $X^{\pi,C}$ and $Y^{\pi,C}$. Therefore, it will be optimal for the agent to do no consumption in the time interval $[0,T)$. The only time where it might be optimal to consume, is the terminal time T. We know that $Y^{\pi,C}$ decreases with time in the absence of consumption, and for small values of $\hat{\mu}$ and $\hat{r}$, $X^{\pi,C}$ is increasing slowly in time. Therefore, $W(X_s^{\pi,C}, Y_s^{\pi,C})$ is decreasing with time from a given starting state (x,y) for sufficiently small values of $\hat{\mu}$ and $\hat{r}$, and it follows that V is increasing in t.

As we will see in Chapter 15, V may also be increasing for some values of t and decreasing for other values of t for fixed x and y. It is difficult to give a general result saying when V is increasing and decreasing, but we will discuss in Chapter 15 which constants that influence whether V is increasing or decreasing in t.

Now we will prove that V is concave in x and y. This is a direct consequence of the concavity of U and W.

Lemma 4.5. *The value function V is concave in x and y on the domain $\overline{\mathscr{D}_T}$.*

Proof. Let $t \in [0,T]$, $(x',y'),(x'',y'') \in \overline{\mathscr{D}}$, $(\pi',C') \in \mathscr{A}_{x',y'}$, $(\pi'',C'') \in \mathscr{A}_{x'',y''}$ and $\lambda \in [0,1]$. Let $X_s'^{\pi',C'}$ and $Y_s'^{\pi',C'}$ be the processes with initial values x' and y', respectively, and let $X_s''^{\pi'',C''}$ and $Y_s''^{\pi'',C''}$ be the processes with initial values x'' and y'', respectively. Define $(x,y) := \lambda(x',y') + (1-\lambda)(x'',y'')$, $C_s := \lambda C_s' + (1-\lambda)C_s''$ and

$$\pi_s := \frac{(\lambda \pi_s' X_s'^{\pi',C'} + (1-\lambda)\pi_s'' X_s''^{\pi'',C''})}{\lambda X_s''^{\pi'',C''} + (1-\lambda) X_s''^{\pi'',C''}},$$

and let $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ denote the processes with initial values x and y, respectively. Using (3.2) and (3.3), we see that

$$X_s^{\pi,C} = \lambda X_s'^{\pi',C'} + (1-\lambda) X_s''^{\pi'',C''}, \tag{4.3}$$

and

$$Y_s^{\pi,C} = \lambda Y_s'^{\pi',C'} + (1-\lambda) Y_s''^{\pi'',C''}. \tag{4.4}$$

Since $X_s^{\pi,C} \geq 0$ for all s and C_s is increasing, we see that $(\pi,C) \in \mathscr{A}_{t,x,y}$. Using (4.3), (4.4) and the concavity of U and W, we get

$$\begin{aligned}
\lambda &\left[\int_t^T e^{-\delta s} U\left(Y_s'^{\pi',C'}\right) ds + W\left(X_T'^{\pi',C'}\right) \right] \\
&\quad + (1-\lambda) \left[\int_t^T e^{-\delta s} U\left(Y_s''^{\pi'',C''}\right) ds + W\left(X_s''^{\pi'',C''}, Y_s''^{\pi'',C''}\right) \right] \\
&= \int_t^T e^{-\delta s} \left[\lambda U\left(Y_s'^{\pi',C'}\right) + (1-\lambda) U\left(Y_s''^{\pi'',C''}\right) \right] ds \\
&\quad + \lambda W\left(X_s'^{\pi',C'}, Y_s'^{\pi',C'}\right) + (1-\lambda) W\left(X_s''^{\pi'',C''}, Y_s''^{\pi'',C''}\right) \\
&\leq \int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W\left(X_s^{\pi,C}, Y_s^{\pi,C}\right),
\end{aligned}$$

which implies

$$\begin{aligned}\lambda \mathrm{E}\Bigg[\int_t^T & e^{-\delta s} U\left(Y_s^{\prime \pi', C'}\right) ds + W\left(X_T^{\prime \pi', C'}, Y_T^{\prime \pi', C'}\right) \Bigg] + \\ & (1-\lambda)\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\prime\prime \pi'', C''}\right) ds + W\left(X_T^{\prime\prime \pi'', C''}, Y_T^{\prime\prime \pi'', C''}\right) \right] \\ & \leq \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi, C}\right) ds + W\left(X_s^{\pi, C}, Y_s^{\pi, C}\right) \right] \\ & \leq V(t, x, y).\end{aligned}$$

Maximizing over $\mathscr{A}_{t,x',y'}$ and $\mathscr{A}_{t,x'',y''}$, we get

$$\lambda V(t, x', y') + (1-\lambda) V(t, x'', y'') \leq V(t, x, y),$$

so V is concave in x and y. □

The following lemma says that V satisfies a similar growth condition as U and W.

Lemma 4.6. *The value function V satisfies $0 \leq V(t,x,y) \leq K(1+x+y)^\gamma$ for all $(t,x,y) \in \overline{\mathscr{D}_T}$ for a constant K independent of t. If U and W satisfy*

$$U(y) \leq K_U y^\gamma \qquad \text{and} \qquad W(x,y) \leq K_W (x+y)^\gamma \tag{4.5}$$

instead of (u_{ii}), we have $0 \leq V(t,x,y) \leq K(x+y)^\gamma$.

Proof. V is obviously non-negative, since the range of both U and W are non-negative.

Suppose $(t,x,y) \in \overline{\mathscr{D}_T}$ for $y > 0$ and $(\pi, C) \in \mathscr{A}_{t,x,y}$, and define a process

$$Z_s = X_s + \frac{Y_s}{\beta},$$

where X_s denotes $X_s^{\pi,C}$, Y_s denotes $Y_s^{\pi,C}$ and Z_s denotes $Z_s^{\pi,C}$. We have $Z_s > 0$, since $X_s \geq 0$ and $Y_s \geq y e^{t-s} > 0$. We also know that Z_s satisfies $Z_t = z$, where $z = x + y/\beta$. Using Itô's formula in Theorem 2.22 and the explicit expressions (3.2) and (3.3) for $X_s^{\pi,C}$ and $Y_s^{\pi,C}$, we get

$$\begin{aligned} dX_s^{\pi,C} = & -dC_s + \big(\hat{r} + (\hat{\mu} - \hat{r})\pi_s\big) X_s^{\pi,C}\, ds + \sigma \pi_s X_s^{\pi,C}\, dB_s \\ & + \pi_{s-} X_{s-}^{\pi,C} \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\, \tilde{N}(ds, dz) \end{aligned}$$

and

$$dY_s^{\pi,C} = -\beta Y_s^{\pi,C}\, ds + \beta\, dC_s.$$

It follows that

$$\begin{aligned} dZ_s = & \big((\hat{r} + (\hat{\mu} - \hat{r})\pi_s) X_s - Y_s\big)\, ds + \sigma \pi_s X_s dB_s \\ & + \pi_{s-} X_{s-} \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\, \tilde{\mu}(ds, dz). \end{aligned}$$

Using Itõ's formula, $\frac{X_{s'}}{Z_{s'}}, \pi_{s'} \frac{X_{s'}}{Z_{s'}} \in [0,1]$, and that the expected value of $\tilde{\mu}$ and B are 0, we get

$$
\begin{aligned}
\mathrm{E}\left[Z_s^\gamma\right] = {} & z^\gamma + \mathrm{E}\Bigg[\int_t^s \gamma Z_{s'}^{\gamma-1}\big((\hat{r}+(\hat{\mu}-\hat{r})\pi_{s'})X_{s'} - Y_{s'}\big)\,ds' \\
& - \frac{1}{2}\gamma(1-\gamma)\int_t^s (\sigma\pi_{s'}X_{s'})^2 Z_{s'}^{\gamma-2}\,ds' \\
& + \int_t^s \int_{\mathbb{R}\setminus\{0\}} \Big(\big(Z_{s'} + \pi_{s'}X_{s'}(e^z-1)\big)^\gamma - Z_{s'}^\gamma \\
& - \gamma\pi_{s'}Z_{s'}^{\gamma-1}X_{s'}(e^z-1)\Big)\,\nu(dz)\,ds'\Bigg] \\
= {} & z^\gamma + \mathrm{E}\Bigg[\int_t^s Z_{s'}^\gamma\Bigg(\gamma(\hat{r}+(\hat{\mu}-\hat{r})\pi_{s'})\frac{X_{s'}}{Z_{s'}} - \gamma\frac{Y_{s'}}{Z_{s'}} \\
& - \frac{1}{2}\gamma(1-\gamma)(\sigma\pi_s')^2\left(\frac{X_{s'}}{Z_{s'}}\right)^2 \\
& + \int_{\mathbb{R}\setminus\{0\}}\left(\left(1+\pi_{s'}\frac{X_{s'}}{Z_{s'}}(e^z-1)\right)^\gamma - 1 - \gamma\pi_{s'}\frac{X_{s'}}{Z_{s'}}(e^z-1)\right)\nu(dz)\Bigg)\,ds'\Bigg] \\
\leq {} & z^\gamma + \mathrm{E}\left[\int_t^s Z_{s'}^\gamma\,ds'\right]k(\gamma).
\end{aligned}
$$

Grönwall's lemma (Theorem 2.28) now gives $\mathrm{E}[Z_s^\gamma] \leq z^\gamma e^{k(\gamma)(s-t)}$, which implies that

$$
\mathrm{E}[X_s^\gamma] \leq K_X(x+y)^\gamma e^{k(\gamma)(s-t)} \tag{4.6}
$$

and

$$
\mathrm{E}[Y_s^\gamma] \leq K_Y(x+y)^\gamma e^{k(\gamma)(s-t)}, \tag{4.7}
$$

where $K_X = \max\{1, \beta^{-\gamma}\}$ and $K_Y = \max\{1, \beta^\gamma\}$. These bounds also hold for $y = 0$ by continuity.

We now get

$$
\begin{aligned}
& \mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y_s\right)\,ds + W\left(X_T, Y_T\right)\right] \\
& \quad \leq \mathrm{E}\left[\int_t^T e^{-\delta s}K_U(1+Y_s)^\gamma\,ds + K_W(1+X_T+Y_T)^\gamma\right] \\
& \quad \leq \mathrm{E}\left[K_U\int_t^T e^{-\delta s}(1+Y_s^\gamma)\,ds + K_W(1+X_T^\gamma+Y_T^\gamma)\right] \\
& \quad \leq K_U\int_t^T e^{-\delta s}\left(1+K_Y(x+y)^\gamma e^{k(\gamma)(s-t)}\right)ds \\
& \qquad + K_W\left(1+(K_X+K_Y)(x+y)^\gamma e^{k(\gamma)(T-t)}\right) \\
& \quad \leq K' + K'(x+y)^\gamma \\
& \quad \leq K(1+x+y)^\gamma,
\end{aligned}
$$

where the constants K' and K can be chosen independently of t, since $t \in [0,T]$ and $[0,T]$ is bounded. Maximizing over $\mathscr{A}_{t,x,y}$, we get

$$
V(t,x,y) \leq K(1+x+y)^\gamma.
$$

If (4.5) holds instead of (u_{ii}), we get

$$V(t,x,y) \le K(x+y)^\gamma$$

for some constant $K \in \mathbb{R}$, because

$$\begin{aligned}
\mathrm{E}&\left[\int_t^T e^{-\delta s} U(Y_s)\, ds + W(X_T, Y_T)\right] \\
&\le \mathrm{E}\left[\int_t^T e^{-\delta s} K_U Y_s^\gamma\, ds + K_W (X_T + Y_T)^\gamma\right] \\
&\le \mathrm{E}\left[K_U \int_t^T e^{-\delta s} Y_s^\gamma\, ds + K_W (X_T^\gamma + Y_T^\gamma)\right] \\
&\le K_U \int_t^T e^{-\delta s} K_Y (x+y)^\gamma e^{k(\gamma)(s-t)}\, ds \\
&\quad + K_W (K_X + K_Y)(x+y)^\gamma e^{k(\gamma)(T-t)} \\
&\le K(x+y)^\gamma.
\end{aligned}$$

□

Using estimates in the proof above, we manage to prove that our value function converges to the value function of the corresponding infinite-horizon problem when $T \to \infty$. The infinite-horizon problem corresponding to our problem, has a value function defined by

$$\widehat{V}(x,y) = \sup_{(\pi,C) \in \widehat{\mathscr{A}_{x,y}}} \mathrm{E}\left[\int_0^\infty e^{-\delta t} U(Y_s^{\pi,C})\, ds\right], \tag{4.8}$$

where

$$\begin{aligned}
X_s^{\pi,C} = x - C_s &+ \int_t^s \left(\hat{r} + (\hat{\mu} - \hat{r})\pi_{s'}\right) X_{s'}^{\pi,C}\, ds' + \int_t^s \sigma \pi_{s'} X_{s'}^{\pi,C}\, dB_{s'} \\
&+ \int_t^s \pi_{s'-} X_{s'-}^{\pi,C} \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\, \tilde{\mu}(ds, dz)
\end{aligned}$$

and

$$Y_s^{\pi,C} = y e^{-\beta(s-t)} + \beta e^{-\beta s} \int_t^s e^{\beta s'}\, dC_{s'}.$$

The initial values of $X^{\pi,C}$ and $Y^{\pi,C}$ are x and y, respectively, and B_t, ν, $\widehat{\mathscr{A}_{x,y}}$ and all the constants involved, satisfy the same conditions as in our problem. We can easily see that the integral term in (3.4) converges to the corresponding integral in (4.8), but we also need to show that the term $W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)$ of (3.4) converges to 0. We need some additional assumptions on W to manage to show this. From a financial point of view, it is reasonable to assume that $W = e^{-\delta T}\overline{W}$ for some function $\overline{W}$ independent of T. This assumption is sufficient to prove that V converges to $\widehat{V}$.

Lemma 4.7. *Let $V^T : \overline{\mathscr{D}_T}$ denote the value function with $T > 0$ as terminal time. Suppose $W = e^{-\delta T}\overline{W}$ for some function $\overline{W}$ independent of T. Then*

$$\lim_{T \to \infty} V^T(0,x,y) = \widehat{V}(x,y)$$

for all $(x,y) \in \overline{\mathscr{D}}$.

Proof. Fix $(x,y) \in \overline{\mathscr{D}}$, let $\{T_n\}_{n\in\mathbb{N}}$ be a sequence converging to ∞, and let $\epsilon > 0$. We will prove the lemma in two steps, first proving

$$V^{T_n}(0,x,y) \geq \widehat{V}(x,y) - \epsilon \tag{4.9}$$

for all sufficiently large n, and then proving

$$V^{T_n}(0,x,y) \leq \widehat{V}(x,y) + \epsilon \tag{4.10}$$

for all sufficiently large n. Let $(\pi, C) \in \widehat{\mathscr{A}_{x,y}}$ be an optimal control for the the infinite-horizon problem, and let $(\pi_n, C_n) \in \mathscr{A}_{0,x,y}$ denote the corresponding control for the problem with terminal time T_n, i.e., $(\pi_n)_s = \pi_s$ and $(C_n)_s = C_s$ for all $s \in [0, T_n]$. The assumption (u_{ii}) in Chapter 3 and the inequality (4.6) in the proof of Lemma 4.6, imply that there is a constant $K \in \mathbb{R}$ such that

$$\begin{aligned} \mathrm{E}\left[\int_{T_n}^{\infty} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds\right] &\leq K_U \mathrm{E}\left[\int_{T_n}^{\infty} e^{-\delta s}\left(1 + Y_s^{\pi,C}\right)^{\gamma}\, ds\right] \\ &\leq K \int_{T_n}^{\infty} e^{-\delta s} \cdot e^{k(\gamma)s}\, ds, \end{aligned}$$

for all $n \in \mathbb{N}$. The constant K may be dependent of x and y, but is independent of n. The right-hand side of this inequality converges to 0 as $n \to \infty$, since $\delta > k(\gamma)$. Therefore there exists an $n \in \mathbb{N}$ such that

$$\mathrm{E}\left[\int_{T_n}^{\infty} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds\right] < \epsilon,$$

for all $n \geq N$. For all $n \geq N$, we get

$$\begin{aligned} V^{T_n}(0,x,y) &\geq \mathrm{E}\left[\int_0^{T_n} e^{-\delta s} U\left(Y_s^{\pi_n,C_n}\right)\, ds + W\left(X_{T_n}^{\pi,C}, Y_{T_n}^{\pi,C}\right)\right] \\ &> \mathrm{E}\left[\int_0^{\infty} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds\right] - \epsilon \\ &= \widehat{V}(x,y) - \epsilon, \end{aligned}$$

and we see that (4.9) holds.

Now we will prove (4.10). Let $(\pi_n, C_n) \in \mathscr{A}_{0,x,y}$ be an optimal control for the system with terminal time T_n. Inequalities (4.6) and (4.7) in the proof of Lemma 4.6 are independent of the control. Using (u_{ii}) from Chapter 3 and inequalities (4.6) and (4.7), we see that there is a constant $K \in \mathbb{R}$, independent of n, such that

$$\begin{aligned} \mathrm{E}\left[W\left(X_{T_n}^{\pi_n,C_n}, Y_{T_n}^{\pi_n,C_n}\right)\right] &= e^{-\delta T_n} \mathrm{E}\left[\overline{W}\left(X_{T_n}^{\pi_n,C_n}, Y_{T_n}^{\pi_n,C_n}\right)\right] \\ &\leq K_W e^{-\delta T_n} \mathrm{E}\left[\left(1 + X_{T_n}^{\pi_n,C_n} + Y_{T_n}^{\pi_n,C_n}\right)^{\gamma}\right] \\ &\leq K e^{-\delta T_n} \cdot e^{k(\gamma) T_n}. \end{aligned}$$

Since $\delta > k(\gamma)$, we see that

$$\mathrm{E}\left[W\left(X_{T_n}^{\pi_n,C_n}, Y_{T_n}^{\pi_n,C_n}\right)\right] \to 0$$

when $n \to \infty$. Let $N \in \mathbb{N}$ be so large that

$$\mathrm{E}\left[W\left(X_{T_n}^{\pi_n,C_n},Y_{T_n}^{\pi_n,C_n}\right)\right]<\epsilon$$

for all $n \geq N$, and let $(\hat{\pi}_n, \widehat{C}_n) \in \widehat{\mathscr{A}_{0,x,y}}$ be defined by

$$(\hat{\pi}_n)_s = \begin{cases} (\pi_n)_s & \text{for } s \leq T_n, \\ 0 & \text{for } s > T_n, \end{cases}$$

and

$$(\widehat{C}_n)_s = \begin{cases} (C_n)_s & \text{for } s \leq T_n, \\ (C_n)_{T_n} & \text{for } s > T_n. \end{cases}$$

We have

$$\begin{aligned} V^{T_n}(0,x,y) = &\mathrm{E}\left[\int_0^{T_n} e^{-\delta s} U\left(Y_s^{\pi_n,C_n}\right) ds + W\left(X_{T_n}^{\pi_n,C_n}, Y_{T_n}^{\pi_n,C_n}\right)\right] \\ \leq &\mathrm{E}\left[\int_0^{\infty} e^{-\delta s} U\left(Y_s^{\hat{\pi}_n,\widehat{C}_n}\right) ds\right] + \epsilon \\ \leq &\widehat{V}(x,y) + \epsilon, \end{aligned}$$

so (4.10) holds. □

Viscosity theory and growth conditions for the value function $\widehat{V}$ are proved in [9]. Using the lemma above and results from [9], we see that the constant K in Lemma 4.6 can be chosen independently of T, provided W is on the form mentioned in Lemma 4.7.

The following lemma says that V is non-decreasing along each line $\beta x + y = C$ for any constant $C > 0$. It will be applied in the continuity proofs in the next section and when proving that V is a viscosity solution of (5.11) in Chapter 6.3. We will also use it in Chapter 7 to derive an optimal consumption strategy.

Lemma 4.8. *If $(t,x',y') \in \overline{\mathscr{D}}$ and $(t,x,y) \in \overline{\mathscr{D}}$ satisfy $x' = x - c$ and $y' = y + \beta c$ for some $c \in (0,x]$, then $V(t,x,y) \geq V(t,x',y')$. We have equality if and only if there is an initial consumption gulp of size $\geq c$ in the optimal consumption plan for (t,x,y).*

Proof. Let the initial time be t, and let the initial values of $X^{\pi,C}$ and $Y^{\pi,C}$ be x and y respectively. If there is an initial consumption gulp of size $c \in (0,x]$, the wealth and average past consumption will become x' and y', respectively. Let $\widehat{\mathscr{A}_{t,x,y}} \subset \mathscr{A}_{t,x,y}$ denote all controls in $\mathscr{A}_{t,x,y}$ that have an initial consumption gulp of size c. We see immediately that $\widehat{\mathscr{A}_{t,x,y}}$ is isomorphic to $\mathscr{A}_{t,x',y'}$, where each control $(\pi,C) \in \widehat{\mathscr{A}_{t,x,y}}$ corresponds to the control $(\pi, C-c) \in \mathscr{A}_{t,x',y'}$. The first part of the lemma follows, because

$$\begin{aligned} V(t,x,y) = &\sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\Big[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\Big] \\ \geq &\sup_{(\pi,C)\in\widehat{\mathscr{A}_{t,x,y}}} \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right] \\ = &\sup_{(\pi',C')\in\mathscr{A}_{t,x',y'}} \mathrm{E}\Big[\int_t^T e^{-\delta s} U\left(Y_s'^{\pi',C'}\right) ds + W\left(X_T'^{\pi',C'}, Y_T'^{\pi',C'}\right)\Big] \\ = &V(t,x',y'), \end{aligned}$$

where $X'^{\pi',C'}$ and $Y'^{\pi',C'}$ denote the processes with initial values x' and y', respectively. We get equality if and only if it was optimal to do an initial consumption gulp of size $\geq c$. □

4.3 Regularity results

The goal of this section is to prove Theorem 4.13, i.e., to prove that V is uniformly continuous on compact subsets of $\overline{\mathscr{D}_T}$. Another interesting result is Theorem 4.10, which says that V is uniformly continuous in x and y for fixed t on the *whole* domain $\overline{\mathscr{D}_T}$, and that the modulus of continuity is independent of t. To prove Theorem 4.13 we show separately that V is continuous in (x, y) and t.

Lemma 4.9 (Uniform continuity in x and y). *The value function V is uniformly continuous in x and y for fixed $t \in [0,T]$, i.e., for each $t \in [0,T]$ there exists a function $\omega_t : \overline{\mathscr{D}} \to (0,\infty)$, such that*

1. *ω_t is continuous at $(0,0)$,*
2. *$\omega_t(0,0) = 0$, and*
3. *ω_t satisfies*
$$V(t,x,y) - V(t,x',y') \leq \omega_t(|x-x'|, |y-y'|)$$
for all $(x,y) \in \overline{\mathscr{D}}$.

Proof. Fix $t \in [0,T]$. We will compare trajectories starting from different points $(x,y),(x',y') \in \overline{\mathscr{D}}$. Let $X^{\pi,C}, Y^{\pi,C}$ denote processes with initial values x, y, and let $X'^{\pi,C}, Y'^{\pi,C}$ denote processes with initial values x', y'. Assume $(\pi, C) \in \mathscr{A}_{t,x,y}$, and define the stopping time

$$\tau = \begin{cases} \inf\{s \in [t,T] : X_s'^{\pi,C} < 0\} & \text{if } X_s'^{\pi,C} < 0 \text{ for some } s \in [t,T], \\ \infty & \text{if } X_s'^{\pi,C} \geq 0 \text{ for all } s \in [t,T]. \end{cases}$$

Define
$$C_s' = C_s \mathbf{1}_{t<\tau} + \big(\Delta X_\tau'^{\pi,C} + X_{\tau^-}'^{\pi,C} + C_\tau\big)\mathbf{1}_{t\geq\tau} \tag{4.11}$$
and
$$\Gamma_s = C_t - C_t'.$$
The processes C_s' and Γ_t are non-decreasing as

$$\begin{aligned} &\Delta X_\tau'^{\pi,C} + X_{\tau^-}'^{\pi,C} + C_\tau \\ &\qquad = \Big(- \Delta C_\tau + \pi_{\tau^-} X_{\tau^-}'^{\pi,C}(e^{\Delta L_\tau} - 1)\Big) + X_{\tau^-}'^{\pi,C} + (C_{\tau^-} + \Delta C_\tau) \\ &\qquad = C_{\tau^-} + X_{\tau^-}'^{\pi,C}(\pi_{\tau^-} e^{\Delta L_\tau} - \pi_{\tau^-} + 1) \\ &\qquad \geq C_{\tau^-}, \end{aligned}$$

where ΔC_τ and ΔL_τ are the values of the control jump and the Lévy process jump, respectively, at time τ. We also see that $X_s'^{\pi,C'} = X_s'^{\pi,C}\mathbf{1}_{s<\tau}$: For $s < \tau$ we obviously have $X_s'^{\pi,C} = X_s'^{\pi,C'}$, as $C_s' = C_s$ for $s < \tau$. We also know that

$$\begin{aligned}
X'^{\pi,C'}_{\tau} &= X'^{\pi,C'}_{\tau^-} + \pi_\tau(e^{\Delta L_\tau}-1)X'^{\pi,C'}_{\tau^-} - \Delta C'^{-}_{\tau}\\
&= X'^{\pi,C'}_{\tau^-} + \pi_\tau(e^{\Delta L_\tau}-1)X'^{\pi,C'}_{\tau^-} - (\Delta X'^{\pi,C}_{\tau} + X'^{\pi,C}_{\tau^-} + C_\tau - C_{\tau^-})\\
&= \pi_\tau(e^{\Delta L_\tau}-1)X'^{\pi,C}_{\tau^-} - \Delta X'^{\pi,C}_{\tau} - \Delta C_\tau\\
&= 0.
\end{aligned}$$

The first equality follows because the change in $X'^{\pi,C'}$ at time τ is a sum of the contributions from the Lévy process and the control. The second equality follows from (4.11). The third equality follows because $C'_{\tau^-} = C_{\tau^-}$ and $X'^{\pi,C}_{\tau^-} = X'^{\pi,C'}_{\tau^-}$. The forth equality follows because the change in $X'^{\pi,C}$ at time τ is a sum of the contributions from the Lévy process and the control.

Since $X'^{\pi,C}_{\tau} = 0$ and C'_s is constant for $s \geq \tau$, we see that $X_s = 0$ for $s \geq \tau$, and $X'^{\pi,C'}_s = X'^{\pi,C}_s \mathbf{1}_{s<\tau}$.

We have $(\pi, C') \in \mathscr{A}_{t,x',y'}$, because $X'^{\pi,C'}_s = X'^{\pi,C}_s \mathbf{1}_{s<\tau} \geq 0$ for all s. We also have $(\pi, \Gamma) \in \mathscr{A}_{|x-x'|,|y-y'|}$. The result obviously holds for $x' \geq x$, because this gives $\Gamma \equiv 0$. It also holds for $x \geq x'$, because $x \geq x'$ implies

$$\begin{aligned}
(X-X')^{\pi,\Gamma}_s &= X^{\pi,C}_s - X'^{\pi,C'}_s\\
&= X^{\pi,C}_s - X'^{\pi,C}_s 1_{s<\tau}\\
&\geq 0.
\end{aligned}$$

We see that

$$|X^{\pi,C}_s - X'^{\pi,C'}_s| = |X-X'|^{\pi,\Gamma}_s,$$

by considering the three cases (1) $x' \geq x$, (2) $x \geq x'$ and $\tau \geq s$ and (3) $x \geq x'$ and $\tau < s$ separately. By the triangle inequality, we also have

$$\begin{aligned}
|Y^{\pi,C}_s - Y'^{\pi,C'}_s| &= \left|(y-y') + \beta e^{-\beta s}\int_t^s e^{\beta s'}\,d\Gamma_{s'}\right|\\
&\leq |(y-y')| + \beta e^{-\beta s}\int_t^s e^{\beta s'}\,d\Gamma_{s'}\\
&= |Y-Y'|^{\pi,\Gamma}_s.
\end{aligned}$$

Using these inequalities, we get

$$\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y^{\pi,C}_s\right)\,ds + W\left(X^{\pi,C}_T, Y^{\pi,C}_T\right)\right]\\
&\quad\leq \mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y'^{\pi,C'}_s\right)\,ds + W\left(X'^{\pi,C}_T, Y'^{\pi,C}_T\right)\right]\\
&\qquad + \mathrm{E}\left[\int_t^T e^{-\delta s}\omega_U(|Y-Y'|^{\pi,\Gamma}_s) + \omega_W(|X-X'|^{\pi,\Gamma}_s, |Y-Y'|^{\pi,\Gamma}_s)\right]\\
&\quad\leq V(t,x',y') + \omega_t(|x-x'|, |y-y'|),
\end{aligned}$$

where ω_U is a modulus of continuity for U, and ω_W is a modulus of continuity for W. The function ω_t is the value function we get when U and W are replaced by ω_U and ω_W, respectively. The utility functions U and W have moduli of continuity, because they are uniformly continuous (see Theorem 2.27), and we know that they are uniformly continuous because they are continuous, concave and non-decreasing (assumption (u_i)

in Chapter 3). The moduli of continuity ω_U and ω_W are assumed to be non-decreasing in x and y.

Maximizing over $\mathscr{A}_{t,x,y}$ and exchanging (x, y) and (x', y') we get

$$|V(t, x, y) - V(t, x', y')| \leq \omega_t(|x - x'|, |y - y'|).$$

The value function ω_t obviously satisfies $\omega_t(0, 0) = 0$. We will now prove that ω_t is continuous at $(0, 0)$. Given any $\epsilon > 0$, there exist constants K_ϵ and L_ϵ such that $\omega_U(y) < \epsilon + K_\epsilon y^\gamma$ and $\omega_W(x, y) < \epsilon + L_\epsilon(x^\gamma + y^\gamma)$. The existence of such constants K_ϵ and L_ϵ follows from assumption (u_{ii}) in Chapter 3. By using (4.6) and (4.7), we get

$$\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s}\omega_U\left(Y_s^{\pi,C}\right)\,ds + \omega_W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right] \\
&\qquad \leq K\epsilon + \mathrm{E}\left[K_\epsilon \int_t^T e^{-\delta s}(Y_s^{\pi,C})^\gamma\,ds\right] + K\epsilon + \mathrm{E}\left[L_\epsilon\Big(\big(X_T^{\pi,C}\big)^\gamma + \big(Y_T^{\pi,C}\big)^\gamma\Big)\right] \\
&\qquad \leq K\epsilon + KK_\epsilon(x + y)^\gamma + K\epsilon + KL_\epsilon(x + y)^\gamma,
\end{aligned} \tag{4.12}$$

for some constant $K > 0$ that is independent of ϵ, x, y, t, π and C. Maximizing over π and C, we get

$$\omega_t(x, y) \leq 2K\epsilon + K(K_\epsilon + L_\epsilon)(x + y)^\gamma.$$

For all x and y sufficiently small, we get

$$\omega_t(x, y) \leq 3K\epsilon. \tag{4.13}$$

Since K is independent of x and y, and ϵ is arbitrary, we see that ω_t is continuous at $(0, 0)$. It follows that ω_t is a modulus of continuity for V for fixed t. □

The next theorem says that the function ω_t in Lemma 4.9 can be chosen independently of t. We will give two alternative proofs of this result. Both proofs are based on the definition of ω_t as the value function when U and W are replaced by their moduli of continuity. The challenging part of the proofs is to show that $\{\omega_t\}_{t\in[0,T]}$ is equicontinuous in t at $(0, 0)$. The first proof shows this directly, while the second proof employs Lemma 4.12 below, which implies that $t \mapsto \omega_t(x, y)$ is continuous for each $(x, y) \in \overline{\mathscr{D}}$.

Theorem 4.10 (Uniform continuity in x and y). *The value function V is uniformly continuous in x and y for fixed $t \in [0, T]$, and the modulus of continuity can be chosen independently of t.*

Proof. We know from Lemma 4.9 that V has a modulus of continuity ω_t for each $t \in [0, T]$, and we see from the proof of the lemma that we may define ω_t by

$$\omega_t = \sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}\omega_U(Y_s^{\pi,C})\,ds + \omega_W(X_T^{\pi,C}, Y_T^{\pi,C})\right],$$

where ω_U and ω_W are moduli of continuity for U and W, respectively. Define $\omega : \overline{\mathscr{D}} \to [0, \infty)$ by

$$\omega(x, y) := \sup_{t\in[0,T]} \omega_t(x, y).$$

We want to prove that ω is a modulus of continuity for V in x and y. We see immediately that $\omega(0,0)=0$ and that

$$|V(t,x,y)-V(t,x',y')| \leq \omega_t(|x-x'|,|y-y'|) \leq \omega(|x-x'|,|y-y'|)$$

for all $t \in [0,T]$ and all $(x,y),(x',y') \in \overline{\mathscr{D}}$. What remains to prove, is that ω is continuous at $(0,0)$.

Given any $\epsilon > 0$, choose K_ϵ and L_ϵ as described in the proof of Lemma 4.9. By using (4.6) and (4.7), we see that the constant K in (4.12) can be chosen independently of t, since the possible values t can take are bounded. It follows that (4.13) holds for all $t \in [0,T]$ for sufficiently small x and y, so

$$\omega(x,y) \leq 3K\epsilon$$

for sufficiently small x and y. We see that ω is continuous at $(0,0)$, and the theorem is proved. □

The next step is to prove that V is continuous in t. This cannot be shown by using techniques in [9], because the authors consider an infinite-horizon problem here. The continuity of V in t may seem obvious, and it is not proved in many articles concerning singular control problems (see for example [32], [13], [25]).

For non-singular problems it is relatively easy to see that regularity in x and y implies regularity in t. If the HJB equation is local, it can be written on the form $v_t + \tilde{F} = 0$, where $\tilde{F} = \tilde{F}(t, X, v, D_X v, D_X^2 v)$. We see that $\tilde{F}$ is well-defined if v is smooth with respect to x and y, and it is reasonable to assume that v_t exists if $\tilde{F}$ is well-defined. We can obtain higher-order regularity in t by differentiating the HJB equation with respect to t, and assuming regularity in x and y. For a singular problem the HJB equation is given by a gradient constraint $G = 0$ on part of the domain, and there we have no information about v_t, even if v is regular in x and y.

The proof below is not based on studying the HJB equation. We will prove continuity by considering the original optimization problem (3.4) and by using Theorem 3.5 (the dynamic programming principle). We also apply Theorem 3.6 (the existence of optimal controls) in the proof below, but we could easily have performed the proof without using this theorem. Instead of using Theorem 3.6, we could have used that, given any $\epsilon > 0$ and $(t,x,y) \in \overline{\mathscr{D}_T}$, there is a $(\pi, C) \in \mathscr{A}_{t,x,y}$ such that

$$V(t,x,y) \leq \mathrm{E}\left[\int_t^\tau e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + V\left(\tau, X_\tau^{\pi,C}, Y_\tau^{\pi,C}\right)\right] + \epsilon$$

for all $\tau \in [t,T]$. This result follows from the dynamic programming principle and the definition of V.

Before we prove the continuity of V in t, we need to prove a technical lemma. We see easily that the limits of the lemma below hold for each $\omega \in \Omega$, but we also need to prove that the *expected value* of the expressions satisfy the limitation results.

Lemma 4.11. *Let $(t,x,y) \in \overline{\mathscr{D}_T}$. If $t < T$, $\{t_n\}_{n\in\mathbb{N}}$ is a sequence such that $t_n \to t^+$, and $(\pi, C) \in \mathscr{A}_{t,x,y}$, we have*

$$\lim_{n\to\infty} \mathrm{E}\left[\int_t^{t_n} e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds\right] = 0. \tag{4.14}$$

If $(\pi^*, C^*) \in \mathscr{A}_{t,x,y}$ *is an optimal control, we also have*

$$\lim_{n\to\infty} \mathrm{E}\left[V\left(t_n, X_{t_n}^{\pi^*,C^*}, Y_{t_n}^{\pi^*,C^*}\right)\right] = \lim_{n\to\infty} \mathrm{E}\big[V(t_n, x, y)\big] \tag{4.15}$$

if either of the limits exist.

If $\{t_n\}_{n\in\mathbb{N}}$ *is a sequence such that* $t_n \to t^-$, *and* $(\pi_n, C_n) \in \mathscr{A}_{t_n,x,y}$ *for all* $n \in \mathbb{N}$, *we have*

$$\lim_{n\to\infty} \mathrm{E}\left[\int_{t_n}^{t} e^{-\delta s} U\left(Y_s^{n,\pi_n,C_n}\right) ds\right] = 0, \tag{4.16}$$

where X^{n,π_n,C_n} *and* Y^{n,π_n,C_n} *denotes the processes with initial values* x *and* y, *respectively, at time* t_n. *We also have*

$$\lim_{n\to\infty} \mathrm{E}\left[V\left(t, X_t^{n,\pi_n,C_n}, Y_t^{n,\pi_n,C_n}\right) - V\left(t, x - \Delta C_n, y + \beta \Delta C_n\right)\right] = 0, \tag{4.17}$$

where $\Delta C_n := (C_n)_t - (C_n)_{t_n^-}$.

Proof. We will divide the proof into four parts, showing (4.14), (4.15), (4.16) and (4.17), respectively.

Part 1: First we assume $\{t_n\}_{n\in\mathbb{N}}$ satisfies $t_n \to t^+$, and that $(\pi, C) \in \mathscr{A}_{t,x,y}$. We want to prove (4.14). We see immediately that the integral of (4.14) converges to 0 for fixed ω, since the length of the interval we integrate over converges to 0, and the integrand is bounded. The challenge is to show that also the expected value of the integral converges to 0, as we do not have any results saying that the convergence of the integral is uniform in ω. Since U is non-negative, it is sufficient to show that, given any $\epsilon > 0$,

$$\mathrm{E}\left[\int_t^{t_n} e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds\right] < \epsilon$$

for sufficiently large n. We define the following set of events $\Omega_m \subset \Omega$ for all $m \in \mathbb{N}$:

$$\Omega_m = \{\omega \in \Omega \,:\, m - 1 \leq C_{\omega,T} < m\}.$$

We see immediately that sets Ω_m are disjoint, and that their union is equal to Ω. The idea of the proof is to split Ω into two sets: One set $\cup_{m=M+1}^{\infty}\Omega_m$ of low probability, consisting of the cases where $Y^{\pi,C}$ grows fast, and one set $\cup_{m=1}^{M}\Omega_m$, consisting of cases where the $Y^{\pi,C}$ satisfies some growth estimate. By Theorem 2.23, we get

$$\mathrm{E}\left[\int_t^{t_n} U(Y_s^{\pi,C})\, ds\right] = \sum_{m=1}^{\infty} \mathrm{E}\left[\int_t^{t_n} U(Y_s^{\pi,C})\, ds \,|\, \Omega_m\right] \mathbb{P}\left[\Omega_m\right].$$

The sum on the right-hand side of this equation is a sum of positive, real numbers. Therefore the following inequality holds for sufficiently large $M \in \mathbb{N}$:

$$\begin{aligned}
\mathrm{E}\bigg[\int_t^{t_n} U(Y_s^{\pi,C})\, ds\bigg] &\leq \sum_{m=1}^{M} \mathrm{E}\left[\int_t^{t_n} U(Y_s^{\pi,C})\, ds \,\middle|\, \Omega_m\right] \mathbb{P}\left[\Omega_m\right] + \epsilon/2 \\
&= \mathrm{E}\left[\int_t^{t_n} U(Y_s^{\pi,C})\, ds \,\middle|\, \cup_{m=1}^{M} \Omega_m\right] \mathbb{P}\left[\cup_{m=1}^{M}\Omega_m\right] + \epsilon/2 \\
&= \mathrm{E}\left[\int_t^{t_n} U(Y_s^{\pi,C})\, ds \,\middle|\, C_T < M\right] \mathbb{P}\left[C_T < M\right] + \epsilon/2.
\end{aligned}$$

The first equality of the above statement, follows from Theorem 2.23 and the fact that the sets Ω_m are disjoint:

$$\begin{aligned}
\mathrm{E}\Big[V(t_n, X_{t_n}, Y_{t_n}) \mid \cup_{m=1}^{M} \Omega_m\Big]\mathbb{P}\left[\cup_{m=1}^{M}\Omega_m\right] \\
+ \mathrm{E}\left[V(t_n, X_{t_n}, Y_{t_n}) \mid \left(\cup_{m=1}^{M}\Omega_m\right)^c\right]\mathbb{P}\left[\left(\cup_{m=1}^{M}\Omega_m\right)^c\right] \\
= \mathrm{E}\big[V(t_n, X_{t_n}, Y_{t_n})\big] \\
= \sum_{m=1}^{\infty}\mathrm{E}\big[V(t_n, X_{t_n}, Y_{t_n}) \mid \Omega_m\big]\mathbb{P}\left[\Omega_m\right] \\
+ \mathrm{E}\left[V(t_n, X_{t_n}, Y_{t_n}) \mid \left(\cup_{m=1}^{M}\Omega_m\right)^c\right]\mathbb{P}\left[\left(\cup_{m=1}^{M}\Omega_m\right)^c\right],
\end{aligned}$$

where ^{c} denotes complement. Since $\mathbb{P}\left[C_T < M\right] \leq 1$ and

$$\mathrm{E}\left[\int_t^{t_n} U(Y_s^{\pi,C})\,ds \,\middle|\, C_T < M\right] < \int_t^{t_n} U(y+\beta M)\,ds \to 0$$

as $n \to \infty$, we see that (4.14) holds.

Part 2: Now suppose $\{t_n\}_{n\in\mathbb{N}}$ satisfies $t_n \to t^+$, and that $(\pi^*, C^*) \in \mathscr{A}_{t,x,y}$ is an optimal control. We will show (4.15) by showing separately that, given any $\epsilon > 0$,

$$\mathrm{E}\left[V(t_n, X_{t_n}^{\pi^*,C^*}, Y_{t_n}^{\pi^*,C^*})\right] \leq \mathrm{E}\big[V(t_n, x, y)\big] + \epsilon \tag{4.18}$$

and

$$\mathrm{E}\left[V(t_n, X_{t_n}^{\pi^*,C^*}, Y_{t_n}^{\pi^*,C^*})\right] > \mathrm{E}\big[V(t_n, x, y)\big] - \epsilon \tag{4.19}$$

for sufficiently large n. We start by showing (4.18). Define the control C_s on $[t,T]$ by

$$C_s = \max\{C_t^*;\, C_s^* - \epsilon\}.$$

Note that C does a jump at time t iff C^* jumps at time t. By the almost certain right-continuity of C^*, C is almost certainly constant on some interval starting at t, before it starts increasing. We have $C_s < C_s^*$ for all $s \in [t,T]$. We also note that $(\pi^*, C) \in \mathscr{A}_{t,x,y}$, because $X_s^{\pi^*,C} \geq X_s^{\pi^*,C^*}$ for all $s \in [t,T]$ and C_s is non-decreasing.

Define $\Omega_m \subset \Omega$ by

$$\Omega_m = \left\{\omega \in \Omega \,:\, m-1 \leq \max\left\{\sup_{s\in(t,T]}\frac{|X_s^{\pi^*,C} - X_t^{\pi^*,C}|}{s-t};\, |C_s^* - C_t^*|\right\} < m\right\}$$

for all $m \in \mathbb{N}$. Note that $\Omega = \Omega_\infty \cup (\cup_{m=1}^{\infty}\Omega_m)$ for some set Ω_∞ of probability 0. We have $\mathbb{P}(\Omega_\infty) = 0$, because C is almost certainly constant on some interval starting with t. Since $\{\Omega_m\}$ are disjoint sets with union Ω, Theorem 2.23 gives us

$$\mathrm{E}\big[V(t_n, X_{t_n}, Y_{t_n})\big] = \sum_{m=1}^{\infty}\mathrm{E}\Big[V(t_n, X_{t_n}, Y_{t_n}) \mid \Omega_m\Big]\mathbb{P}(\Omega_m).$$

The sum on the right-hand side of this equation consists of positive real numbers, and therefore, given any $\epsilon > 0$, there is an $M \in \mathbb{N}$ such that

$$\begin{aligned}\mathrm{E}\Big[V(t_n,X_{t_n},Y_{t_n})\Big] &< \sum_{m=1}^{M}\mathrm{E}\Big[V(t_n,X_{t_n},Y_{t_n})\,|\,\Omega_m\Big]\mathbb{P}(\Omega_m)+\epsilon\\ &=\mathrm{E}\Big[V(t_n,X_{t_n},Y_{t_n})\,|\,\cup_{m=1}^{M}\Omega_m\Big]\mathbb{P}\left(\cup_{m=1}^{M}\Omega_m\right)+\epsilon.\end{aligned}$$

As for Part 1, the equality of the above statement follows from Theorem 2.23. If $\omega\in\cup_{m=1}^{M}\Omega_m$, we know that $X_{t_n}\le X_t+M(t_n-t)$ and $Y_{t_n}\le Y_t+\beta M(t_n-t)$. Letting $\hat{\omega}$ be a modulus of continuity for V in x and y, we get

$$\begin{aligned}&\mathrm{E}\Big[V(t_n,X_{t_n},Y_{t_n})\,|\,\cup_{m=1}^{M}\Omega_m\Big]\mathbb{P}\left(\cup_{m=1}^{M}\Omega_m\right)\\ &\qquad\le V\big(t_n,X_t+M(t_n-t),Y_t+\beta M(t_n-t)\big)\\ &\qquad\le V(t_n,X_t,Y_t)+\hat{\omega}\big(M(t_n-t),\beta M(t_n-t)\big).\end{aligned}$$

Since $\hat{\omega}(0,0)=0$ and $\hat{\omega}$ is continuous at $(0,0)$, we see that (4.18) holds.

Now we want to prove (4.19). Since $(\pi^*,C^*)\in\mathscr{A}_{t,x,y}$ is an optimal control, we have

$$\begin{aligned}&\mathrm{E}\left[\int_t^{t_n}e^{-\delta s}U(Y_s^{\pi^*,C^*})\,ds+V\left(t_n,X_{t_n}^{\pi^*,C^*},Y_{t_n}^{\pi^*,C^*}\right)\right]\\ &\quad=V(t,x,y)\\ &\quad\ge\mathrm{E}\left[\int_t^{t_n}e^{-\delta s}U\left(Y_s^{0,0}\right)\,ds+V\left(t_n,X_{t_n}^{0,0},Y_{t_n}^{0,0}\right)\right].\end{aligned}\tag{4.20}$$

We know that

$$V\left(t_n,X_{t_n}^{0,0},Y_{t_n}^{0,0}\right)=V\left(t_n,xe^{\hat{r}(t_n-t)},ye^{-\beta(t_n-t)}\right).$$

By the result in Part 1, we know that both integral terms in (4.20) converge to 0 as $n\to\infty$. Since V is continuous in x and y, we see that (4.19) holds.

Part 3: Now assume $\{t_n\}_{n\in\mathbb{N}}$ is a sequence such that $t_n\to t^-$. Define Ω_m by

$$\Omega_m=\begin{cases}\{\omega\in\Omega\,|\,\sup_{n\in\mathbb{N}}(C_n)_t\in[m-1,m)\} & \text{if } m\in\mathbb{N},\\ \{\omega\in\Omega\,|\,\sup_{n\in\mathbb{N}}(C_n)_t=\infty\} & \text{if } m=\infty\end{cases}$$

for all $m\in\mathbb{N}\cup\{\infty\}$. Note that $\mathbb{P}(\Omega_\infty)=0$, and that the sets Ω_m are disjoint. Theorem 2.23 gives us

$$\mathrm{E}\left[\int_{t_n}^{t}e^{-\delta s}U(Y_s^{\pi,C})\,ds\right]=\sum_{m=1}^{\infty}\mathrm{E}\left[\int_{t_n}^{t}e^{-\delta s}U(Y_s^{\pi,C})\,ds\,\Big|\,\Omega_m\right]\mathbb{P}[\Omega_m].$$

Proceeding as in Parts 1 and 2, we see that, given any $\epsilon>0$, there is an $M\in\mathbb{N}$, such that

$$\begin{aligned}\mathrm{E}\left[\int_{t_n}^{t}e^{-\delta s}U(Y_s^{\pi,C})\,ds\right] &< \sum_{m=1}^{M}\mathrm{E}\left[\int_{t_n}^{t}e^{-\delta s}U(Y_s^{\pi,C})\,ds\,\Big|\,\Omega_m\right]\mathbb{P}[\Omega_m]+\epsilon\\ &=\mathrm{E}\left[\int_{t_n}^{t}e^{-\delta s}U(Y_s^{\pi,C})\,ds\,\Big|\,\cup_{m=1}^{M}\Omega_m\right]\mathbb{P}\left[\cup_{m=1}^{M}\Omega_m\right]+\epsilon.\end{aligned}$$

We have

$$\mathrm{E}\left[\int_{t_n}^{t}e^{-\delta s}U(Y_s^{\pi,C})\,ds\,\Big|\,\cup_{m=1}^{M}\Omega_m\right]<\int_{t_n}^{t}e^{-\delta s}U(y+\beta M)\,ds\to 0$$

as $n \to \infty$, so (4.16) holds.

Part 4: Let $\hat{\omega}$ be a modulus of continuity for V in x and y, i.e.,

$$\begin{aligned}\mathrm{E}\Big[\big|V\left(t, X_t^{n,\pi_n,C_n}, Y_t^{n,\pi_n,C_n}\right) - V\big(t, x-\Delta C_n, y+\beta\Delta C_n\big)\big|\Big] \\ \leq \mathrm{E}\Big[\hat{\omega}\big(\big|X_t^{n,\pi_n,C_n} - x + \Delta C_n\big|, \big|Y_t^{n,\pi_n,C_n} - y - \beta\Delta C_n\big|\big)\Big]\end{aligned}$$

for all $n \in \mathbb{N}$, see Theorem 4.10. We wish to show that the right-hand side of this inequality is less than ϵ for sufficiently large n, where $\epsilon > 0$ is some given real number. Define $\Omega_m \subset \Omega$ by

$$\Omega_m = \left\{\omega \in \Omega \,\middle|\, m-1 \leq \max\left\{\sup_{n\in\mathbb{N}} \frac{|X_t^{n,\pi_n,C_n} + \Delta C_n - x|}{(t-t_n)^{1/2}}; \sup_{n\in\mathbb{N}} \Delta C_n\right\} < m\right\}$$

for all $m \in \mathbb{N}$. Note that the sets Ω_m are disjoint, and that $\Omega = \Omega_\infty \cup (\cup_{m=1}^\infty \Omega_m)$ for some set Ω_∞ of probability 0. Using Theorem 2.23, we get

$$\begin{aligned}&\mathrm{E}\left[\hat{\omega}\left(\left|X_t^{n,\pi_n,C_n} + \Delta C_n - x\right|, \left|Y_t^{n,\pi_n,C_n} - y - \beta\Delta C_n\right|\right)\right] \\ &\quad = \sum_{m=1}^{\infty} \mathrm{E}\left[\hat{\omega}\left(\left|X_t^{n,\pi_n,C_n} + \Delta C_n - x\right|, \left|Y_t^{n,\pi_n,C_n} - y - \beta\Delta C_n\right|\right) \mid \Omega_m\right] \mathbb{P}(\Omega_m).\end{aligned}$$

Let $\epsilon > 0$. Proceeding as in Parts 1-3, we see that there is an $M \in \mathbb{N}$, such that

$$\begin{aligned}&\mathrm{E}\left[\hat{\omega}\left(\left|X_t^{n,\pi_n,C_n} + \Delta C_n - x\right|, \left|Y_t^{n,\pi_n,C_n} - y - \beta\Delta C_n\right|\right)\right] \\ &\quad < \sum_{m=1}^{M} \mathrm{E}\left[\hat{\omega}\left(\left|X_t^{n,\pi_n,C_n} + \Delta C_n - x\right|, \left|Y_t^{n,\pi_n,C_n} - y - \beta\Delta C_n\right|\right) \mid \Omega_m\right] \mathbb{P}(\Omega_m) + \epsilon \\ &\quad = \mathrm{E}\left[\hat{\omega}\left(\left|X_t^{n,\pi_n,C_n} + \Delta C_n - x\right|, \left|Y_t^{n,\pi_n,C_n} - y - \beta\Delta C_n\right|\right) \mid \cup_{m=1}^M \Omega_m\right] \mathbb{P}(\cup_{m=1}^M \Omega_m) \\ &\qquad + \epsilon \\ &\quad < \hat{\omega}\Big((t-t_n)^{1/2} M,\, y\big|(e^{-\beta(t-t_n)} - 1)\big| + \beta e^{-\beta t}\big|e^{-\beta t_n} - e^{-\beta t}\big| M\Big) + \epsilon.\end{aligned}$$

The right-hand side of this inequality converges to 0 as $n \to \infty$, and (4.17) follows. □

Using the results of the lemma above, we manage to prove continuity of V in t.

Lemma 4.12. *The value function V is continuous in t.*

Proof. Let $(t,x,y) \in \overline{\mathscr{D}_T}$. We want to show that V is right-continuous and left-continuous at (t,x,y).

We will start by proving that V is right-continuous for $t < T$. Let $\{t_n\}_{n\in\mathbb{N}}$ be a sequence in $(t,T]$ such that $t_n \to t^+$ as $n \to \infty$. By Theorems 3.5 and 3.6, we have

$$V(t,x,y) = \mathrm{E}\left[\int_t^{t_n} e^{-\delta s} U\left(Y_s^{\pi^*,C^*}\right) ds + V\left(t_n, X_{t_n}^{\pi^*,C^*}, Y_{t_n}^{\pi^*,C^*}\right)\right] \tag{4.21}$$

for all $n \in \mathbb{N}$, where $(\pi^*, C^*) \in \mathscr{A}_{t,x,y}$ is an optimal control. The first term on the right-hand side of (4.21) converges to 0 by Lemma 4.11. It follows that $V\left(t_n, X_{t_n}^{\pi^*,C^*}, Y_{t_n}^{\pi^*,C^*}\right)$ converges, and by Lemma 4.11, we see that $\lim_{n\to\infty} V(t_n,x,y)$ exists, and that the right-hand side of (4.21) converges to $\lim_{n\to\infty} V(t_n, x, y)$. It follows that V is right-continuous.

Now we will show that V is left-continuous at (t, x, y). Let $\{t_n\}_{n\in\mathbb{N}}$ be a sequence in $[0, t)$ such that $t_n \to t^-$ as $n \to \infty$. Define $\Delta t_n = t - t_n$. First we will prove that, for all $\epsilon > 0$,

$$V(t_n, x, y) \geq V(t, x, y) - \epsilon \tag{4.22}$$

for all sufficiently large n. Then we will prove that, for all $\epsilon > 0$,

$$V(t_n, x, y) \leq V(t, x, y) + \epsilon \tag{4.23}$$

for all sufficiently large n. These two results will imply that V is left-continuous at t.

For any $(\pi, C) \in \mathscr{A}_{t_n,x,y}$, let $X^{n,\pi,C}$ and $Y^{n,\pi,C}$ denote processes with initial values x and y, respectively, at time t_n. Assume $(\pi^*, C^*) \in \mathscr{A}_{t,x,y}$ is an optimal control, and define $(\pi_n, C_n) \in \mathscr{A}_{t_n,x,y}$ by

$$(\pi_n)_s = \begin{cases} \pi^*_{n,s+\Delta t_n} \text{ if } s \leq T - \Delta t_n, \\ 0 \text{ if } s > T - \Delta t_n \end{cases}$$

and

$$(C_n)_s = \begin{cases} C^*_{s+\Delta t_n} \text{ if } s \leq T - \Delta t_n, \\ C^*_T \text{ if } s > T - \Delta t_n. \end{cases}$$

If we let the initial time of the Lévy process be t_n for each $n \in \mathbb{N}$, we see that $X_s^{n,\pi_n,C_n} = X_{s+\Delta t_n}^{\pi^*,C^*}$ and $Y_s^{n,\pi_n,C_n} = Y_{s+\Delta t_n}^{\pi^*,C^*}$ for all $n \in \mathbb{N}$ and all $s \in [t, T - \Delta t_n]$. We also see that

$$X_s^{n,\pi_n,C_n} = X_{T-\Delta t_n}^{n,\pi_n,C_n} + \int_{T-\Delta t_n}^{s} \hat{r} X_{s'}^{n,\pi_n,C_n}\, ds' = X_{T-\Delta t}^{n,\pi_n,C_n} e^{\hat{r}\left(s-(T-\Delta t_n)\right)}$$

and

$$Y_s^{n,\pi_n,C_n} = Y_{T-\Delta t_n}^{n,\pi_n,C_n} e^{-\beta\left(s-(T-\Delta t_n)\right)}$$

for all $s \in [T - \Delta t_n, T]$. Using these expressions, we get

$$\begin{aligned}
V(t_n, x, y) \geq{} & \mathrm{E}\left[\int_{t_n}^{T} e^{-\delta s} U\left(X_s^{n,\pi_n,C_n}\right) ds + W\left(X_T^{n,\pi_n,C_n}, Y_T^{n,\pi_n,C_n}\right)\right] \\
={} & \mathrm{E}\left[\int_{t_n}^{T-\Delta t_n} e^{-\delta s} U\left(Y_s^{n,\pi_n,C_n}\right) ds + W\left(X_{T-\Delta t_n}^{n,\pi_n,C_n}, Y_{T-\Delta t_n}^{n,\pi_n,C_n}\right)\right] \\
& + \mathrm{E}\Bigg[\int_{T-\Delta t_n}^{T} e^{-\delta s} U\left(Y_s^{n,\pi_n,C_n}\right) ds + W\left(X_T^{n,\pi_n,C_n}, Y_T^{n,\pi_n,C_n}\right) \\
& - W\left(X_{T-\Delta t_n}^{n,\pi_n,C_n}, Y_{T-\Delta t_n}^{n,\pi,C}\right)\Bigg] \\
={} & V(t, x, y) + \mathrm{E}\Bigg[\int_{T-\Delta t_n}^{T} e^{-\delta s} U\left(Y_{T-\Delta t_n}^{n,\pi_n,C_n} e^{-\beta(s-T+\Delta t_n)}\right) ds \\
& + W\left(X_{T-\Delta t_n}^{n,\pi_n,C_n} e^{\hat{r}\Delta t_n}, Y_{T-\Delta t_n}^{n,\pi_n,C_n} e^{-\beta \Delta t_n}\right) \\
& - W\left(X_{T-\Delta t_n}^{n,\pi_n,C_n}, Y_{T-\Delta t_n}^{n,\pi_n,C_n}\right)\Bigg] \\
={} & V(t, x, y) + \mathrm{E}\Bigg[\int_{T-\Delta t_n}^{T} e^{-\delta s} U\left(Y_T^{\pi^*,C^*} e^{-\beta(s-T+\Delta t_n)}\right) ds
\end{aligned}$$

$$+ W\left(X_T^{\pi^*,C^*} e^{\hat{r}\Delta t_n}, Y_T^{\pi^*,C^*} e^{-\beta \Delta t_n}\right) - W\left(X_T^{\pi^*,C^*}, Y_T^{\pi^*,C^*}\right)\Bigg]. \tag{4.24}$$

Equation (4.22) is proved if we can show that the right-hand side of this equation converges to $V(t,x,y)$ as $n \to \infty$. The second term on the right-hand side of (4.24) (the term containing an expected value), is decreasing as n increases, because it is decreasing for each $\omega \in \Omega$. The second term on the right-hand side of (4.24) is therefore convergent. It cannot converge to something positive, as it converges to 0 for each $\omega \in \Omega$, and it cannot converge to something negative, as it is positive for all $n \in \mathbb{N}$ and all $\omega \in \Omega$. It must therefore converge to 0, and we see that (4.22) holds.

Now we will prove (4.23). By Theorems 3.5 and 3.6, there is a $(\pi_n^*, C_n^*) \in \mathscr{A}_{t,x,y}$ such that

$$V(t_n, x, y) = \mathrm{E}\left[\int_{t_n}^t e^{-\delta s} U\left(Y_s^{n,\pi_n^*,C_n^*}\right) ds + V\left(t, X_t^{n,\pi_n^*,C_n^*}, Y_t^{n,\pi_n^*,C_n^*}\right)\right]$$

for each $n \in \mathbb{N}$. Note that

$$\lim_{n\to\infty} \mathrm{E}\left[\int_{t_n}^t e^{-\delta s} U\left(Y_s^{n,\pi_n^*,C_n^*}\right) ds\right] = 0$$

by Lemma 4.11, so what remains to be proved is that, for each $\epsilon > 0$,

$$\mathrm{E}\left[V\left(t, X_t^{n,\pi_n^*,C_n^*}, Y_t^{n,\pi_n^*,C_n^*}\right)\right] \leq V(t,x,y) + \epsilon \tag{4.25}$$

for sufficiently large $n \in \mathbb{N}$. We have

$$\begin{aligned}
\mathrm{E}&\Big[V\Big(t, X_t^{n,\pi_n^*,C_n^*}, Y_t^{n,\pi_n^*,C_n^*}\Big) - V(t,x,y)\Big] \\
&= \mathrm{E}\left[V\left(t, X_t^{n,\pi_n^*,C_n^*}, Y_t^{n,\pi_n^*,C_n^*}\right) - V\left(t, X_t^{n,\pi_n^*,C_n^*}, Y_t^{n,\pi_n^*,C_n^*}\right)\right] \\
&\quad + \mathrm{E}\left[V\left(t, X_t^{n,\pi_n^*,C_n^*}, Y_t^{n,\pi_n^*,C_n^*}\right) - V(t,x,y)\right],
\end{aligned}$$

where $\Delta C_n := (C_n)_t - (C_n)_{t_n^-}$. The first term of this equation converges to 0 by Lemma 4.11, and the second term converges to something ≤ 0 by Lemma 4.8, and (4.25) follows. □

We are now ready to prove the main theorem of the section.

Theorem 4.13. *The value function V is uniformly continuous on compact subsets of $\overline{\mathscr{D}_T}$.*

Proof. Let $\mathscr{O}_T$ be a compact subset of $\overline{\mathscr{D}_T}$. Lemmas 4.9 and 4.12 imply that V is continuous, and by the Heine-Cantor Theorem 2.26, V is uniformly continuous on $\mathscr{O}_T$. □

Before we end the section, we will give an alternative proof of Theorem 4.10. The alternative proof of Theorem 4.10 is based on the following technical lemma.

Lemma 4.14. *Let $f : A \times B \to \mathbb{R}$ be continuous, $B \subset \mathbb{R}^m$ be compact, and $A \subset \mathbb{R}^n$, where $m, n \in \mathbb{N}$. For each $b \in B$, the function $a \mapsto f(a,b)$ is uniformly continuous in a, and $a \mapsto f(a,b)$ has a modulus of continuity ω_b that is continuous in a and b. Then it is possible to choose a modulus of continuity for $a \mapsto f(a,b)$ that is independent of b.*

Proof. Define $\omega : (\mathbb{R}^+)^n \to \mathbb{R}$ by

$$\omega(a) := \sup_{b \in B} \omega_b(a),$$

where $\mathbb{R}^+ = [0, \infty)$. We want to prove that ω is a modulus of continuity for $a \mapsto f(a, b)$ for all $b \in B$. We see immediately that

$$|f(a, b) - f(a', b)| \leq \omega(|a - a'|),$$

for all $a, a' \in A$ and $b \in B$, so what remains to be proved is that ω is well-defined, ω is continuous at (0), and that $\omega(0) = 0$.

The function ω is well-defined if it is finite everywhere. Fix $a \in A$. The function $b \mapsto \omega_b(a)$ is defined on a compact set, and therefore it must obtain its maximum value for some $b^* \in B$. We know that $\omega_{b^*}(a)$ is finite, and therefore $\omega(a)$ is also finite.

We obviously have $\omega(0) = 0$, as $\omega_b = 0$ for all $b \in B$, so what remains to be proved, is that ω is continuous at 0. This will be proved by contradiction. Assume $\{a_n\}_{n \in \mathbb{N}}$ is a sequence in A that converges to 0, and that $\omega(a_n) > \epsilon$ for all $n \in \mathbb{N}$ and some $\epsilon > 0$. Since B is compact, we have $\omega(a_n) = \omega_{b_n}(a_n)$ for some $b_n \in B$. We can assume without loss of generality that $\{b_n\}_{n \in \mathbb{N}}$ is convergent; if $\{b_n\}_{n \in \mathbb{N}}$ was not convergent, we could obtained a convergent sequence by taking a subsequence. Assume $b_n \to b^*$ as $n \to \infty$. We see that $(b_n, a_n) \to (b^*, 0)$ as $n \to \infty$, and by the continuity of ω_b in a and b, we know that $\omega_{b_n}(a_n) \to \omega_{b^*}(0) = 0$ as $n \to \infty$. This is a contradiction to the assumption that $\omega_{b_n}(a_n) > \epsilon$ for all $n \in \mathbb{N}$, and we see that ω is continuous at 0. □

We now give the second proof of Theorem 4.10.

Proof of Theorem 4.10: This proof is based on the lemma above. In the proof of Lemma 4.9, we saw that ω_t is a modulus of continuity for $(x, y) \mapsto V(t, x, y)$, where ω_t is the value function we get by replacing U and W by ω_U and ω_W, respectively. By Lemmas 4.9 and 4.12, we see that ω_t is continuous in (x, y) and t. Since we know that $[0, T]$ is compact, Lemma 4.14 applies, and we see that ω_t in Lemma 4.9 can be chosen independently of t. □

Chapter 5
The Hamilton-Jacobi-Bellman Equation of the optimization problem

In this chapter the Hamilton-Jacobi-Bellman (HJB) equation corresponding to the control problem will be derived. We will also prove that the HJB equation's dimension can be reduced in the case of CRRA utility.

Note that the derivation of the HJB equation is not a proof, rather a justification. The technique is inspired by [31], where the authors derive an HJB equation for another singular stochastic control problem. The main difference between our problem and the problem considered in [31], is that the value S_t of the risky asset in [31] follows a geometric Brownian motion, where the drift and volatility vary with t and S_t. In our problem the risky asset is a geometric Lévy process with constant drift and volatility. The jump part of the Lévy process results in a new non-local term in the HJB equation, and having constant drift and volatility reduces the dimension of the HJB equation by one compared to in [31]. It is possible to derive the HJB equation using other methods, see for example [25].

Fix $(t,x,y) \in \mathscr{D}_T$, and let $(\pi, C) \in \mathscr{A}_{t,x,y}$ be a control. We want to prove that the value function V, defined by (3.4), satisfies a specific differential equation, the HJB equation, at (t,x,y). The following assumptions are made in this chapter:

- (v_i) V is twice continuously differentiable on $\overline{\mathscr{D}_T}$, and its derivative and second derivative are bounded.
- (v_{ii}) If an optimal control C is discontinuous at time t, the size of the consumption gulp is a differentiable function of x and y.
- (v_{iii}) C_s is either discontinuous at t, or it is almost certainly differentiable at t and continuous on an interval $[t, t+\Delta t]$. There is a constant $c > 0$, such that we have either $C'(t) = 0$ or $C'(t) > c$ for differentiable C.

That V is twice continuously differentiable is not a realistic assumption, as we know that the value function of many singular control problems is not even differentiable. We make the assumptions in order to manage to derive the HJB equation using dynamic programming.

Using Itô's formula from Theorem 2.22 and the explicit expressions (3.2) and (3.3) for $X_s^{\pi,C}$ and $Y_s^{\pi,C}$, we get

$$\begin{aligned} dX_s^{\pi,C} =& -dC_s + (\hat{r} + (\hat{\mu} - r)\pi_s)X_s^{\pi,C}\,ds + \sigma\pi_s X_s^{\pi,C}\,dB_s \\ &+ \pi_{s-}X_{s-}^{\pi,C}\int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\,\tilde{N}(ds,dz) \end{aligned} \tag{5.1}$$

and

$$dY_s^{\pi,C} = -\beta Y_s^{\pi,C}\,ds + \beta\,dC_s. \tag{5.2}$$

Since $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ are semimartingales, and V is twice continuously differentiable by assumption (v_i), Itô's formula gives us

$$\begin{aligned}
&V(t{+}\Delta t, X_{t+\Delta t}^{\pi,C}, Y_{t+\Delta t}^{\pi,C})\\
&\quad= V(t, X_{t^-}^{\pi,C}, Y_{t^-}^{\pi,C}) + \int_t^{t+\Delta t} \frac{\partial V_s}{\partial t} + \frac{\partial V_s}{\partial x}\hat{r}(1-\pi_s)X_s^{\pi,C} + \frac{\partial V_s}{\partial x}\pi_s\hat{\mu}X_s^{\pi,C}\\
&\qquad - \frac{\partial V_s}{\partial y}\beta Y_s^{\pi,C} + \frac{1}{2}\frac{\partial^2 V_s}{\partial x^2}(\sigma\pi_s X_s^{\pi,C})^2\,ds\\
&\qquad + \int_t^{t+\Delta t} \frac{\partial V_s}{\partial x}\sigma\pi_s X_s^{\pi,C}\,dB_s + \int_{t^-}^{t+\Delta t} -\frac{\partial V_{s^-}}{\partial x} + \beta\frac{\partial V_{s^-}}{\partial y}\,dC_s\\
&\qquad + \int_{t^-}^{t+\Delta t} X_{s^-}^{\pi,C}\pi_{s^-}\frac{\partial V_{s^-}}{\partial x}\int_{\mathbb{R}\setminus\{0\}}(e^z-1)\,\tilde{N}(ds,dz)\\
&\qquad + \sum_{t\le s\le t+\Delta t}\left(\Delta V_s - \frac{\partial V_{s^-}}{\partial x}\Delta X_s^{\pi,C} - \frac{\partial V_{s^-}}{\partial y}\Delta Y_s^{\pi,C}\right).
\end{aligned} \tag{5.3}$$

We let V_s denote $V\left(s, X_s^{\pi,C}, Y_s^{\pi,C}\right)$ in this equation, (5.4) and (5.5), while partial derivatives are described using the ∂ symbol. This notation will only be used in (5.3)-(5.5); in the other equations subscripts mean partial derivatives.

The stochastic variables B and $\tilde{N}$ are martingales, and therefore their expectation is equal to 0. Using this, the dynamic programming principle of Theorem 3.5 and equation (5.3), we get

$$\begin{aligned}
0 = \sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\Bigg[&\int_t^{t+\Delta t} e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds\\
&+ \int_t^{t+\Delta t}\left(\frac{\partial V_s}{\partial t} + \frac{\partial V_s}{\partial x}\hat{r}(1-\pi_s)X_s^{\pi,C} + \frac{\partial V_s}{\partial x}\pi_s\hat{\mu}X_s^{\pi,C} - \frac{\partial V_s}{\partial y}\beta Y_s^{\pi,C}\right.\\
&\left.+ \frac{1}{2}\frac{\partial^2 V_s}{\partial x^2}(\sigma\pi_s X_s^{\pi,C})^2\right)ds + \int_{t^-}^{t+\Delta t}\left(-\frac{\partial V_{s^-}}{\partial x} + \beta\frac{\partial V_{s^-}}{\partial x}\right)dC_s\\
&+ \sum_{t\le s\le t+\Delta t}\left(\Delta V_s - \frac{\partial V_{s^-}}{\partial x}\Delta X_s^{\pi,C} - \frac{\partial V_{s^-}}{\partial y}\Delta Y_s^{\pi,C}\right)\Bigg].
\end{aligned} \tag{5.4}$$

If C is continuous in the interval $[t, t+\Delta]$,

$$\mathrm{E}\left(\sum_{t\le s\le t+\Delta t}\Delta V_s - \frac{\partial V_{s^-}}{\partial x}\Delta X_s^{\pi,C} - \frac{\partial V_{s^-}}{\partial y}\Delta Y_s^{\pi,C}\right) = \int_t^{t+\Delta t}\mathscr{I}^{\pi_s}\left(s, X_s^{\pi,C}, Y_s^{\pi,C}, V\right)\,ds, \tag{5.5}$$

where

$$\begin{aligned}
\mathscr{I}^{\pi}(t,x,y,V) := \int_{\mathbb{R}\setminus\{0\}} &V\big(t, x+\pi x(e^z-1), y\big) - V(t,x,y)\\
&- \pi x(e^z-1)V_x(t,x,y)\,\nu(dz).
\end{aligned}$$

We can derive this equality by using that

$$\Delta Y_s^{\pi,C} = 0,$$

$$\Delta X_s^{\pi,C} = \pi_{s^-} X_{s^-}^{\pi,C}(e^{\Delta L} - 1)$$

and

$$\Delta V_s = V\left(s, X_{s^-}^{\pi,C} + \Delta X_s^{\pi,C}, Y_{s^-}^{\pi,C}\right) - V\left(s, X_{s^-}^{\pi,C}, Y_{s^-}^{\pi,C}\right)$$

if C is continuous, and that ν is a measure such that $\nu(A)$ is the expected number of jumps per unit time of L whose size belongs to A. The integral defining $\mathscr{I}^\pi$ is finite by Lemma 3.4.

We set $dC \equiv 0$ in (5.4), divide by Δt, let $\Delta t \to 0$, use (5.5), assumption (v_i) and that $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ are right-continuous, and get

$$\begin{aligned}0 \geq\, & e^{-\delta t}U(y) + V_t - \beta y V_y \\ & + \sup_{\pi \in [0,1]}\left[V_x \hat{r}(1-\pi)x + V_x \pi \hat{\mu} x + \frac{1}{2}V_{xx}(\sigma\pi x)^2 + \mathscr{I}^\pi(t,x,y,V)\right].\end{aligned} \tag{5.6}$$

Let $(\hat{\pi}, \widehat{C}) \in \mathscr{A}_{t,x,y}$ be an optimal control. The existence of an optimal control follows from Theorem 3.6. We consider three different cases:

(1) there is a consumption gulp at time t,
(2) $\widehat{C}$ is continuous on an interval $[t, t+\Delta t]$ and $\widehat{C}'(t) > c$, and
(3) $\widehat{C}$ is continuous on an interval $[t, t+\Delta t]$ and $\widehat{C}'(t) = 0$.

There may exist $\omega \in \Omega$ not covered by any of these cases, but by assumption (v_{iii}), we will not consider these cases here.

We start with case (1). By Lemma 4.8,

$$V(t,x,y) = V(t, x - \Delta C, y + \beta \Delta C) \tag{5.7}$$

for $\Delta C = \Delta\widehat{C} := \widehat{C}_t - \widehat{C}_{t^-}$. Since $(\hat{\pi}, \widehat{C})$ is an optimal control, $\Delta\widehat{C}$ is the value of $\Delta\widehat{C}$ that maximizes the right-hand side of (5.7). Therefore, the derivative of the right-hand side of (5.7) with respect to ΔC must be 0 at $\Delta\widehat{C}$, i.e.,

$$-V_x(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C}) + \beta V_y(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C}) = 0. \tag{5.8}$$

Differentiating (5.7) with respect to x, using assumption (v_{ii}) and applying equation (5.8), we get

$$\begin{aligned}V_x(t,x,y) =& \Big(-V_x(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C}) + \beta V_y(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C})\Big)\frac{\partial \Delta\widehat{C}}{\partial x} \\ & + V_x(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C}) \\ =& V_x(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C}).\end{aligned} \tag{5.9}$$

Similarly we get $V_y(t,x,y) = V_y(t, x - \Delta\widehat{C}, y + \beta\Delta\widehat{C})$. Combining (5.8), (5.9) and this equation, we get

$$V_x(t,x,y) = \beta V_y(t,x,y).$$

Now we consider case (2), and let $c > 0$ be such that $\widehat{C}'(t) > c$. We know that $(\hat{\pi}, \widehat{C})$ maximizes the right-hand side of (5.4), because it is an optimal control. We make a small perturbation from this control to (π, C), where $C_s = \widehat{C}_s + C_s^\Delta$ and $\pi_s \equiv \hat{\pi}_s$ for some function $C_s^\Delta : [t,T] \to \mathbb{R}$ defined by $C_s^\Delta := c'(s-t)$, $c' \in \mathbb{R}$. It is sufficient to define (π, C) on $[t, t+\Delta t]$. Using (5.1) and (5.2), we see that $X_s^{\pi,C} = X_s^{\hat\pi,\widehat{C}} - C_s^\Delta + O(\Delta t^2)$ and $Y_s^{\pi,C} = Y_s^{\hat\pi,\widehat{C}} + \beta C_s^\Delta$. We have not yet checked that $(\pi, C) \in \mathscr{A}_{t,x,y}$, but will do so below. We insert C into (5.4), and make an approximation to this equation, assuming Δt is small.

We have

$$\begin{aligned}
&\int_t^{t+\Delta t} f(s, X_s^{\pi,C}, Y_s^{\pi,C})\,ds - \int_t^{t+\Delta t} f(s, X_s^{\hat\pi,\widehat{C}}, Y_s^{\hat\pi,\widehat{C}})\,ds \\
&\qquad \approx \int_t^{t+\Delta t} -C_s^\Delta f_x\left(s, X_s^{\hat\pi,\widehat{C}}, Y_s^{\hat\pi,\widehat{C}}\right) + \beta C_s^\Delta f_y\left(s, X_s^{\hat\pi,\widehat{C}}, Y_s^{\hat\pi,\widehat{C}}\right)\,ds \\
&\qquad = O(\Delta t^2)
\end{aligned}$$

and

$$\begin{aligned}
&\int_t^{t+\Delta t} f(s, X_s^{\pi,C}, Y_s^{\pi,C})\,dC_s - \int_t^{t+\Delta t} f(s, X_s^{\hat\pi,\widehat{C}}, Y_s^{\hat\pi,\widehat{C}})\,d\widehat{C}_s \\
&\qquad \approx c' \int_t^{t+\Delta t} f(s, X_s^{\pi,C}, Y_s^{\pi,C})\,ds - \int_t^{t+\Delta t} C_s^\Delta f_x(s, X_s^{\hat\pi,\widehat{C}}, Y_s^{\hat\pi,\widehat{C}})\,d\widehat{C}_s \\
&\qquad\quad + \int_t^{t+\Delta t} \beta C_s^\Delta f_y(s, X_s^{\hat\pi,\widehat{C}}, Y_s^{\hat\pi,\widehat{C}})\,d\widehat{C}_s \\
&\qquad = c' \int_t^{t+\Delta t} f(s, X_s^{\pi,C}, Y_s^{\pi,C})\,ds + O(\Delta t^2)
\end{aligned}$$

for any sufficiently smooth function f. The equation (5.5) is valid, since the consumption is continuous on $[t, t+\Delta t]$. Replacing all functions in (5.4) with the approximations stated above, and using the fact that C^* maximizes the right-hand side of (5.4), we see that

$$\mathrm{E}\left[c' \int_t^{t+\Delta t} -V_x + \beta V_y\,ds\right] \le 0. \tag{5.10}$$

We have $(\pi, C) \in \mathscr{A}_{t,x,y}$ on $[t, t+\Delta]$ for all $c' \in (-c, 0]$ and sufficiently small Δt, because C is increasing and $X_s^{\pi,C} \ge X_s^{\hat\pi,\widehat{C}} \ge 0$ for $c' \in (-c, 0]$. Now assume $c' > 0$, i.e., $C_s > \widehat{C}_s$ for $s \in [t, t+\Delta]$. We see that C_s is strictly increasing, so $(\pi, C) \in \mathscr{A}_{t,x,y}$ iff $X_s^{\pi,C} \ge 0$ on $[t, t+\Delta t]$. Since $x > 0$, we will have $(\pi, C) \in \mathscr{A}_{t,x,y}$ for sufficiently small Δt. We see that (5.10) holds for all $c' \in (-c, \infty)$ when $\Delta t \to 0$, i.e., for both positive and negative values of c'. By continuity, we must have $-V_x + \beta V_y = 0$.

Finally we consider case (3). Since $(\hat\pi, \widehat{C})$ is maximizing the right-hand side of (5.4), we have $\beta V_y - V_x \le 0$ by the same argument as in case (2), but with only positive values of c' valid, as C should be non-decreasing. Equation (5.5) is valid, since the consumption is continuous. We insert (5.5) and $(\pi, C) = (\hat\pi, \widehat{C})$ into (5.4). All terms in (5.4) are of order $O(\Delta t)$, except for $\int_{t^-}^{t+\Delta t} -V_x + \beta V_y\,dC$, which is of order $O(C_{t+\Delta t} - C_t)$. We know that $C_{t+\Delta t} - C_t$ is of smaller order than Δt, since $C'(t) = 0$ for all $c > 0$. By letting $\Delta t \to 0$ in (5.4), we see that (5.6) holds as an equality.

Summing up we have showed that V satisfies

$$0 = \max\left\{G(D_X v);\; v_t + F\big(t, X, D_X v, D_X^2 v, \mathscr{I}^\pi(t, X, v)\big)\right\}, \tag{5.11}$$

where

$$\begin{aligned} F\big(t, X, &D_X v, D_X^2 v, \mathscr{I}^\pi(t,X,v)\big) = e^{-\delta t}U(y) - \beta y v_y(t,X) \\ &+ \max_{\pi\in[0,1]} \left[(\hat{r}(1-\pi)+\pi\hat{\mu})x v_x(t,X) + \frac{1}{2}(\sigma\pi x)^2 v_{xx}(t,X) + \mathscr{I}^\pi(t,X,v) \right] \end{aligned} \tag{5.12}$$

and

$$G(D_X v) = \beta v_y(t,X) - v_x(t,X). \tag{5.13}$$

See Section 3.1 for the definition of $D_X v$ and $D_X^2 v$. We see from the above derivation that $G(D_X v) = 0$ is associated with consumption, while

$$v_t + F\big(t, X, D_X v, D_X^2 v, \mathscr{I}^\pi(t,X,v)\big) = 0$$

is associated with no consumption. The equation (5.11) is the *Hamilton-Jacobi-Bellman (HJB) equation* of our optimization problem. Under assumptions $(v_i) - (v_{iii})$ it is valid for all $(t, X) \in \mathscr{D}_T$. Note that the upper-case function V denotes the value function, while we will let the lower-case function v denote a general function defined on $\overline{\mathscr{D}_T}$.

5.1 Reduction of the dimension

In this section we show that the dimension of the HJB equation can be reduced if there is a $\gamma \in (0,1)$ such that U and W satisfy

$$U(\alpha y) = \alpha^\gamma U(y),\ W(\alpha x, \alpha y) = \alpha^\gamma W(x,y) \quad \forall \alpha \geq 0,\ (x,y) \in \overline{\mathscr{D}}. \tag{5.14}$$

By inserting $y = 1$ in (5.14), we see that U can be written on the form $U(y) = y^\gamma/\gamma$ (possibly multiplied by a constant), i.e., it is a CRRA utility function, see Section 2.3.

A reduction of the dimension will be a large advantage when solving the equation numerically, as the computing time and need of memory will decrease. We will transform (5.11) to a one-dimensional equation in three different ways. The reason for doing it in several different ways, is that the three HJB equations we get exhibit different numerical properties, see Chapter 13.

If (5.14) holds, we see directly from (3.4) that

$$V(t,x,y) = y^\gamma \overline{V}\left(t, \frac{x}{y}\right)$$

for some function $\overline{V} : [0,T] \times [0,\infty) \to [0,\infty)$. Inserting this relation into equation (5.11), we obtain the following HJB equation:

$$\max\left\{ y^{\gamma-1}\overline{G}(r,\overline{v},\overline{v}_r);\ y^\gamma\left(\overline{v}_t + \overline{F}\big(t,r,\overline{v},\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^\pi(t,r,\overline{v})\big)\right)\right\} = 0, \tag{5.15}$$

where

$$\overline{G}(r,\overline{v},\overline{v}_r) = -\beta\overline{v}_r r + \beta\gamma\overline{v} - \overline{v}_r, \tag{5.16}$$

$$\begin{aligned} \overline{F}\big(t,r,\overline{v},\overline{v}_r,&\overline{v}_{rr},\mathscr{I}^\pi(t,r,\overline{v})\big) = U(1)e^{-\delta t} + \beta r\overline{v}_r - \beta\gamma\overline{v} \\ &+ \max_{\pi\in[0,1]}\left[(\hat{r}(1-\pi)+\pi\hat{\mu})r\overline{v}_r + \frac{1}{2}(\sigma\pi r)^2\overline{v}_{rr} + \overline{\mathscr{I}}^\pi(t,r,\overline{v})\right] \end{aligned} \tag{5.17}$$

and

$$\overline{\mathscr{I}}^{\pi}(t,r,\overline{v}) = \int_{\mathbb{R}\setminus 0} \overline{v}\big(t, r+\pi r(e^z-1)\big) - \overline{v}(t,r) - \pi r \overline{v}_r(t,r)(e^z-1)\,\nu(dz).$$

Equation (5.15) is equivalent to

$$\max\Big\{\overline{G}(\overline{v},\overline{v}_r);\ \overline{v}_t + \overline{F}\big(t,r,\overline{v},\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\big)\Big\} = 0.$$

By (4.1) and (4.2), we have

$$\overline{V}(t,0) = \int_t^T e^{-\delta s} U\left(e^{-\beta(s-t)}\right) ds + W\left(0, e^{-\beta(T-t)}\right) \tag{5.18}$$

and

$$\overline{V}(T,r) = W(r,1) \tag{5.19}$$

for all $t \in [0,T]$ and all $r \geq 0$.

If we define $r = x/(x+y)$ instead, we get

$$V(t,x,y) = (x+y)^{\gamma}\overline{V}(t,r)$$

for some function $\overline{V}$. Inserting into (5.11), we get the following HJB equation:

$$\max\Big\{(x+y)^{\gamma-1}\overline{G}(r,\overline{v},\,\overline{v}_r),(x+y)^{\gamma}\Big(\overline{v_t} + \overline{F}\big(t,r,\overline{v},\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\big)\Big)\Big\},$$

where

$$\overline{G}(r,\overline{v},\overline{v}_r) = \gamma(\beta-1)\overline{v} - (1-r+\beta r)\overline{v}_r, \tag{5.20}$$

$$\begin{aligned} \overline{F}\big(t,r,\overline{v},\overline{v}_r,\overline{v}_{rr},\mathscr{I}^{\pi}(t,r,\overline{v})\big) &= e^{-\delta t}U(1-r) - \beta\left(\gamma(1-r)\overline{v} + (1-r)^2\overline{v}_r\right) \\ &+ \max_{\pi\in[0,1]}\Big[\big(\hat{r}(1-\pi)+\pi\hat{\mu}\big)\big(\gamma r\overline{v} + r(1-r)\overline{v}_r\big) + \frac{1}{2}(\sigma\pi)^2\Big[\gamma(\gamma-1)r^2\overline{v} \\ &+ 2(\gamma-1)r^2(1-r)\overline{v}_r + \overline{v}_{rr}r^2(1-r)^2\Big] + \overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\Big] \end{aligned} \tag{5.21}$$

and

$$\mathscr{I}^{\pi}(t,r,\overline{v}) = \int_{\mathbb{R}\setminus\{0\}} \overline{v}\left(t, \frac{(1+\pi(e^z-1))r}{r\pi(e^z-1)+1}\right) - (1+r\gamma)\overline{v}(t,r) - r(1-r)\overline{v}_r\,\nu(dz).$$

By (4.1) and (4.2), we have

$$\overline{V}(T,r) = W(r,1-r)$$

and

$$\overline{V}(t,0) = \frac{U(1)}{\delta+\beta\gamma}\left(e^{-\delta t} - e^{-(\delta+\beta\gamma)T+\beta\gamma t}\right).$$

If we define $r = y/x$, we have

$$V(t,x,y) = x^{\gamma}\overline{V}(t,r)$$

for some function $\overline{V}$, and the HJB equation becomes

$$\max\Big\{x^{\gamma-1}\overline{G}(r,\overline{v},\overline{v}_r);\ x^{\gamma}\Big(v_t + \overline{F}\big(t,r,\overline{v},\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\big)\Big)\Big\},$$

where

$$\overline{G}(r,\overline{v},\overline{v}_r) = \beta\overline{v}_r + r\overline{v}_r - \gamma\overline{v}, \tag{5.22}$$

$$\begin{aligned}\overline{F}\big(t,r,\overline{v},\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\big) &= e^{-\delta t}U(r) - \beta r\overline{v}_r \\ &+ \max_{\pi\in[0,1]}\Bigg[\Big(\hat{r}(1-\pi)+\pi\hat{\mu}\Big)(\gamma\overline{v}-r\overline{v}_r) \\ &+ \frac{1}{2}(\sigma\pi)^2\Big(\gamma(\gamma-1)\overline{v}+2r(1-\gamma)\overline{v}_r+r^2\overline{v}_{rr}\Big)+\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\Bigg]\end{aligned} \tag{5.23}$$

and

$$\begin{aligned}\overline{\mathscr{I}}^{\pi}(t,r,\overline{v}) = \int_{\mathbb{R}\setminus\{0\}} &\overline{v}\left(t,\frac{r}{1+\pi(e^z-1)}\right)-\overline{v}(t,r) \\ &-\pi(e^z-1)\Big(\gamma\overline{v}(t,r)-r\overline{v}_r(t,r)\Big)\,\nu(dz).\end{aligned}$$

We have

$$\overline{V}(T,r)=W(1,r)$$

for all $r\in\mathbb{R}$ and

$$\overline{V}(t,r)=\frac{r^{\gamma}U(1)e^{\beta\gamma t}}{\delta+\beta\gamma}\left(e^{-(\delta+\beta\gamma)t}-e^{-(\delta+\beta\gamma)T}\right)+r^{\gamma}e^{-\beta\gamma(T-t)}W(0,1)$$

for large r and all $t\in[0,T]$.

If we assume

$$U(y)=\ln y,\;\; W(x,y)=\ln x, \tag{5.24}$$

the dimension of the problem can also be reduced. Note that both U and W are CRRA utility functions in this case. We see from (3.4) that

$$V(\alpha x,\alpha y,t)=V(t,x,y)+g(t)\ln\alpha,$$

where

$$g(t)=1+\frac{1}{\delta}\left(e^{-\delta t}-e^{-\delta T}\right),$$

which implies that

$$V(t,x,y)=\overline{V}(t,r)+g(t)\ln y$$

for $r=x/y$. The HJB equation becomes

$$\max\Big\{\overline{G}(t,r,v_r);\;\overline{v}_t+\overline{F}\big(t,r,\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\big)\Big\}=0,$$

where

$$\begin{aligned}\overline{G}(t,r,v_r) &= -\beta\overline{v}_r r+g(t)\beta-\overline{v}_r,\\ \overline{F}\big(t,r,\overline{v}_r,\overline{v}_{rr},\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\big) &= \beta\big(\overline{v}_r r-g(t)\big)\\ &+\max_{\pi\in[0,1]}\left[\big(\hat{r}(1-\pi)+\pi\hat{\mu}\big)r\overline{v}_r+\frac{1}{2}(\sigma\pi r)^2\overline{v}_{rr}+\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})\right]\end{aligned}$$

and

$$\overline{\mathscr{I}}^{\pi}(t,r,\overline{v})=\int_{\mathbb{R}\setminus\{0\}}\overline{v}\big(t,r+r\pi(e^z-1)\big)-\overline{v}(t,r)-r\pi(e^z-1)\overline{v}_r(t,r)\,\nu(dz).$$

However, note that U and W satisfying (5.24) are not consistent with the assumptions we did in Chapter 3, because they take negative values. Therefore much of the theory developed in the thesis is not valid for this problem.

Chapter 6
Viscosity theory for the Hamilton-Jacobi-Bellman equation

In this chapter we develop viscosity solution theory for the terminal value problem (5.11) and (4.2). The goal of the chapter is to prove that the value function V defined by (3.4), is the unique constrained viscosity solution of (5.11) and (4.2).

Definitions and some initial results are given in Section 6.1. We show that the notion of viscosity solutions can be formulated equivalently in several different ways, and show two lemmas that connect the theory of viscosity solutions and the theory of classical solutions. It is advantageous to have several equivalent definitions of viscosity solutions, as some theorems are easier to prove with one definition than the others.

In Section 6.2 some results needed to prove a comparison principle are given. The proof of the comparison principle is more difficult than in the case of a local HJB operator, because it is more challenging to find a formulation of viscosity solutions by use of so-called subjets and superjets. In addition to preparing results needed for the comparison principle proof, the section contains a lemma concerning the existence of some test functions that previously have been used in a subjet/superjet definition of viscosity solutions.

Existence and uniqueness results are shown in Section 6.3. The uniqueness proof is based on a comparison principle, which we will prove using results from Section 6.2.

Some proofs in Sections 6.2 and 6.3 are based on techniques found in [8], [9] and [35], where the authors consider an infinite-horizon version of our problem, but the time-dependency of the value function in our problem requires new approaches in some cases. Some proofs of convergence are also given in more detail here than in [8] and [9].

6.1 Definitions and first results

In this section we will first define what we mean by classical solutions and viscosity solutions. Then we will prove three lemmas that give alternative ways of defining viscosity solutions, and at the end of the section we will give two lemmas that connect the theory of classical solutions and viscosity solutions

Definition 6.1 (Classical solutions).

(1) Let $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$. Any function $v \in C^{1,2,1}(\overline{\mathscr{D}_T})$ is a classical subsolution (supersolution) of (5.11) *in $\mathscr{O}_T$ if and only if*

$$\max\big\{G(D_Xv);\, v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big)\big\} \geq 0\,(\leq 0) \tag{6.1}$$

for all $(t,X) \in \mathscr{O}_T$.

(2) Any $v \in C^{1,2,1}(\overline{\mathscr{D}_T})$ *is a classical solution of* (5.11) *if and only if*

$$\max\big\{G(D_Xv);\, v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big)\big\} = 0$$

for all $(t,X) \in \overline{\mathscr{D}_T}$.

(3) Any $v \in C^{1,2,1}(\overline{\mathscr{D}_T})$ *is a classical solution the terminal value problem* (5.11) *and* (4.2) *if and only if* v *is a classical solution of* (5.11) *and* v *satisfies* (4.2).

If the value function V is smooth, one can show by verification theorems that it satisfies (5.11) in a classical sense, see for example [25] and [31]. However, the value function is often not smooth enough to be a solution of the HJB equation in a classical sense, and therefore one has introduced the *viscosity solution*.

Definition 6.2 (Viscosity solutions).

(1) Let $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$. *Any function* $v \in C(\overline{\mathscr{D}_T})$ *is a viscosity subsolution (supersolution) of* (5.11) *in* $\mathscr{O}_T$ *if and only if we have, for every* $(t,X) \in \mathscr{O}_T$ *and* $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ *such that* (t,X) *is a global maximum (minimum) relative to* $\mathscr{O}_T$ *of* $v - \phi$,

$$\max\big\{G(D_X\phi);\, \phi_t + F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\big\} \geq 0\,(\leq 0). \tag{6.2}$$

(2) Any $v \in C(\overline{\mathscr{D}_T})$ *is a constrained viscosity solution of* (5.11) *if and only if* v *is a viscosity supersolution of* (5.11) *in* $\mathscr{D}_T$, *and* v *is a viscosity subsolution of* (5.11) *in* $[0,T) \times \overline{\mathscr{D}}$.

(3) Any $v \in C(\overline{\mathscr{D}_T})$ *is a constrained viscosity solution of the terminal value problem* (5.11) *and* (4.2) *if and only if* v *is a viscosity solution of* (5.11) *and* v *satisfies* (4.2).

Note that a viscosity solution of (5.11) satisfies the supersolution equation only on the domain $\mathscr{D}_T$, while the subsolution equation must be satisfied on the whole domain $[0,T) \times \overline{\mathscr{D}}$. The subsolution equation serves as a boundary condition for $x = 0$ and $y = 0$.

Note also that the explicit solution formula (4.1) for $x = 0$ satisfies the viscosity subsolution equation (6.2): We know nothing about the sign of G, but by putting $x = 0$ and (4.1) into $V_t + F$, we observe that

$$V_t(t,0,y) + F\big(t,0,y,D_XV,D_X^2V,\mathscr{I}^\pi(t,0,y,V)\big) = V_t(t,0,y) + e^{-\delta t}U(y) - \beta y V_y(t,0,y) = 0,$$

so (6.2) is satisfied. We will also see in other parts of the thesis that

$$V_t(t,X) + F\big(t,X,D_XV,D_X^2V,\mathscr{I}^\pi(t,X,V)\big) \geq 0$$

is associated with the boundary $x = 0$.

Before we prove the three lemmas that give alternative formulations of viscosity solutions, we need to prove a technical lemma, which will be useful in the proof of several theorems in this and later chapters. Note that the lemma also may be used to show that

$$\mathscr{I}^\pi(t_n, X_n, \phi) \to \mathscr{I}^\pi(t, X, \phi)$$

when $(t_n, X_n) \to (t, X)$, and that

$$\mathscr{I}^\pi(t, X, \phi_n) \to \mathscr{I}^\pi(t, X, \phi)$$

when $\phi_n \to \phi$, $(\phi_n)_x \to \phi_x$ and $(\phi_n)_{xx} \to \phi_{xx}$.

Lemma 6.3. *Let $\phi, \phi_n \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t_n, X_n), (t, X) \in \overline{\mathscr{D}_T}$ for all $n \in \mathbb{N}$. Assume $(t_n, X_n) \to (t, X)$, $\phi_n(s, Y) \to \phi(s, Y)$, $(\phi_n)_x(s, Y) \to \phi_x(s, Y)$ and $(\phi_n)_{xx}(s, Y) \to \phi_{xx}(s, Y)$ for all $(s, Y) \in \overline{\mathscr{D}_T}$. Then*

$$\mathscr{I}^\pi(t_n, X_n, \phi_n) \to \mathscr{I}^\pi(t, X, \phi)$$

as $n \to \infty$.

Proof. We have

$$\begin{aligned} |\mathscr{I}^\pi(t_n, &X_n, \phi_n) - \mathscr{I}^\pi(t, X, \phi)| \\ &\leq |\mathscr{I}^\pi(t_n, X_n, \phi_n) - \mathscr{I}^\pi(t_n, X_n, \phi)| + |\mathscr{I}^\pi(t_n, X_n, \phi) - \mathscr{I}^\pi(t, X, \phi)|, \end{aligned} \tag{6.3}$$

so it is sufficient to prove that the two terms on the right-hand side of (6.3) converge to 0.

We start to consider the first term on the right-hand side. We have

$$\begin{aligned} &\left|\mathscr{I}^\pi(t_n, X_n, \phi_n) - \mathscr{I}^\pi(t_n, X_n, \phi)\right| \\ &\quad = \left| \int_{\mathbb{R}\setminus\{0\}} (\phi_n - \phi)\big(t_n, x_n + x_n\pi(e^z - 1), y_n\big) - (\phi_n - \phi)(t_n, X_n) \right. \\ &\quad\quad \left. - x_n\pi(e^z - 1)(\phi_n - \phi)_x(t_n, X_n)\, \nu(dz) \right| \\ &\quad \leq \left| \int_{|z|\leq 1} (\phi_n - \phi)\big(t_n, x_n + x_n\pi(e^z - 1), y_n\big) - (\phi_n - \phi)(t_n, X_n) \right. \\ &\quad\quad \left. - x_n\pi(e^z - 1)(\phi_n - \phi)_x(t_n, X_n)\, \nu(dz) \right| \\ &\quad\quad + \left| \int_{1<|z|\leq R} (\phi_n - \phi)\big(t_n, x_n + x_n\pi(e^z - 1), y_n\big) - (\phi_n - \phi)(t_n, X_n) \right. \\ &\quad\quad \left. - x_n\pi(e^z - 1)(\phi_n - \phi)_x(t_n, X_n)\, \nu(dz) \right| \\ &\quad\quad + \left| \int_{|z|>R} (\phi_n - \phi)\big(t_n, x_n + x_n\pi(e^z - 1), y_n\big) - (\phi_n - \phi)(t_n, X_n) \right. \\ &\quad\quad \left. - x_n\pi(e^z - 1)(\phi_n - \phi)_x(t_n, X_n)\, \nu(dz) \right| \end{aligned} \tag{6.4}$$

for all $R > 1$. The first term on the right-hand side of (6.4) converges to 0 when $n \to \infty$: By applying (3.1), doing a Taylor expansion of $\phi_n - \phi$ around (t_n, X_n), and using $(e^z-1)^2 < 3z^2$ for $|z|<1$ and the Heine-Cantor theorem (Theorem 2.26), we see that

$$\begin{aligned}
&\bigg|\int_{|z|\leq 1} (\phi_n-\phi)(t_n, x_n + x_n\pi(e^z-1), y_n) - (\phi_n-\phi)(t_n, X_n) \\
&\qquad - x_n\pi(e^z-1)(\phi_n-\phi)_x(t_n,X_n)\,\nu(dz)\bigg| \\
&\quad \leq 3(\pi x_n)^2 \sup_{|z|\leq 1}\big|(\phi_n-\phi)_{xx}(t_n, x_n + x_n\pi(e^z-1), y_n)\big| \int_{|z|\leq 1} z^2\,\nu(dz) \\
&\quad \to 0.
\end{aligned}$$

The second term on the right-hand side of (6.4) converges to 0 for all values of R by (3.1), and because $\{z\in\mathbb{R} : 1\leq |z|\leq R\}$ is compact:

$$\begin{aligned}
&\bigg|\int_{1<|z|\leq R} (\phi_n-\phi)\big(t_n, x_n + x_n\pi(e^z-1), y_n\big) - (\phi_n-\phi)(t_n, X_n) \\
&\qquad - x_n\pi(e^z-1)(\phi_n-\phi)_x(t_n,X_n)\,\nu(dz)\bigg| \\
&\quad \leq 2 \sup_{1<|z|\leq R}\Big(\big|(\phi_n-\phi)(t_n, x_n + x_n\pi(e^z-1), y_n)\big| \\
&\qquad + \big|x_n\pi(e^z-1)(\phi_n-\phi)_x(t_n,X_n)\big|\Big)\int_{1<|z|\leq R}\nu(dz) \\
&\quad \to 0.
\end{aligned}$$

The last term on the right-hand side of (6.4) converges to 0 when we let $R\to\infty$, and the convergence is uniform in n: $(\phi_n-\phi)_x(t_n,X_n)\to 0$ as $n\to\infty$ by the Heine-Cantor Theorem. Defining $\psi = A(1+x+y)$, $A\in\mathbb{R}$, and using $\phi,\phi_n\in C_1(\overline{\mathscr{D}_T})$, we see that $|\phi_n-\phi|\leq\psi$ for all $n\in\mathbb{N}$ for some sufficiently large value of A. This implies that there exists a constant $B\in\mathbb{R}$ independent of n such that

$$\begin{aligned}
&\big|(\phi_n-\phi)(t_n, x_n + x_n\pi(e^z-1), y_n) - (\phi_n-\phi)(t_n,X_n)\big| \\
&\quad \leq \psi(t_n, x_n + x_n\pi(e^z-1), y_n) - \psi(t_n,X_n) + B \\
&\quad = Ax_n\pi(e^z-1) + B.
\end{aligned}$$

We know that $Ax_n\pi(e^z-1)+B$ is integrable with respect to ν, and we see that the integrand of the last term in (6.4) is bounded by some integrable function that is independent of n. It follows that the last term on the right-hand side of (6.4) converges to 0 as $R\to\infty$, and that the convergence is uniform in n.

We conclude that the first term on the right-hand side of (6.3) converges to 0 as $n\to\infty$.

Now we want to prove that the second term on the right-hand side of (6.3) converges to 0 as $n\to\infty$. By the middle value theorem,

$$\begin{aligned}
&\mathscr{I}^\pi(t_n, X_n, \phi) - \mathscr{I}^\pi(t, X, \phi) \\
&\quad = \int_{\mathbb{R}\setminus\{0\}} I_t(t_{z,n}, X_{z,n}, z)(t_{\epsilon_n}-t) + I_x(t_{z,n}, X_{z,n}, z)(x_{\epsilon_n}-x) \qquad (6.5)\\
&\qquad + I_y(t_{z,n}, X_{z,n}, z)(y_{\epsilon_n}-y)\,\nu(dz),
\end{aligned}$$

where $I : \overline{\mathscr{D}_T} \times \mathbb{R} \to \mathbb{R}$ denotes the integrand of $\mathscr{I}^\pi$, and $(t_{z,n}, X_{z,n})$ is some point on the line between (t_n, X_n) and (t, X). We see that the integral on the right-hand side is well-defined, because the integrand is equal to the difference between two integrable functions.

The derivative of I with respect to x is given by

$$\begin{aligned} I_x(t_{z,n}, X_{z,n}, z) = & v_x(t_{z,n}, x_{z,n} + \pi(e^z - 1)x_{z,n}, y_{z,n})\big(1 + \pi(e^z - 1)\big) \\ & - v_x(t_{z,n}, X_{z,n}) - \pi(e^z - 1)v_{xx}(t_{z,n}, X_{z,n}). \end{aligned}$$

By a Taylor expansion, we see that I_x is of order $|e^z - 1|$ for large $|z|$, so I_x is bounded by some integrable function for large $|z|$. See the proof of Lemma 3.4 for a proof of this. For small z we see by a Taylor expansion that I_x is of order z^2, so also in this case I_x is bounded by some integrable function. Note that we do not claim that I_x itself is integrable; all we have proved is that the integrand on the right-hand side of (6.5) is integrable, and that I_x is bounded by some function that is integrable with respect to ν.

We get similar results for I_y and I_t. Assume I_x, I_y and I_t are all bounded by the function $f : \overline{\mathscr{D}_T} \to \mathbb{R}$, and that f is integrable with respect to ν. It follows from the dominated convergence theorem (Theorem 2.8) that the right-hand side of (6.5) converges to 0 as $n \to \infty$. □

Before a second formulation of viscosity solutions is given, we will introduce two new integral operators: For any $\kappa \in (0, 1)$, $(t, X) \in \overline{\mathscr{D}_T}$, $\phi \in C_1(\overline{\mathscr{D}_T}) \cap C^{1,2,1}(\overline{\mathscr{D}_T})$ and $P = (p_1, p_2) \in \mathbb{R}^2$, define

$$\begin{aligned} \mathscr{I}^{\pi,\kappa}(t, X, \phi, P) := \int_{|z|>\kappa} \Big(& \phi(t, x + \pi x(e^z - 1), y) \\ & - \phi(t, X) - \pi x p_1(e^z - 1)\Big)\, \nu(dz) \end{aligned}$$

and

$$\begin{aligned} \mathscr{I}^\pi_\kappa(t, X, \phi) := \int_{|z|\le\kappa} \Big(& \phi(t, x + \pi x(e^z - 1), y) - \phi(t, X) \\ & - \pi x_1(e^z - 1)\phi_x(t, X)\Big)\, \nu(dz). \end{aligned}$$

Both integrals are well-defined by a similar argument as in the proof of Lemma 3.4. Note that

$$\mathscr{I}^\pi(t, X, \phi) = \mathscr{I}^{\pi,\kappa}(t, X, \phi, D_X\phi) + \mathscr{I}^\pi_\kappa(t, X, \phi),$$

and that

$$\lim_{\kappa\to 0} \mathscr{I}^\pi_\kappa(t, X, \phi) = 0.$$

We define

$$\begin{aligned} & F\big(t,X, D_X\phi, D^2_X\phi, \mathscr{I}^{\pi,\kappa}(t, X, v, D_X\phi), \mathscr{I}^\pi_\kappa(t, X, \phi)\big) \\ & \quad := F\big(t, X, D_X\phi, D^2_X\phi, \mathscr{I}^{\pi,\kappa}(t, X, v, D_X\phi) + \mathscr{I}^\pi_\kappa(t, X, \phi)\big) \end{aligned}$$

for all $(t, X) \in \overline{\mathscr{D}_T}$, $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $v \in C_1(\overline{\mathscr{D}_T})$.

Now we give a formulation of viscosity solutions that is equivalent to Definition 6.2 for $v \in C_1(\overline{\mathscr{D}_T})$. The definition given in Definition 6.2 can be used to prove existence of viscosity solutions, while a formulation using the new integral operators $\mathscr{I}^{\pi,\kappa}$ and $\mathscr{I}^\pi_\kappa$ is more appropriate for proving uniqueness and comparison results.

Theorem 6.4. *Let $v \in C_1(\overline{\mathscr{D}_T})$. Then v is a subsolution (supersolution) of* (5.11) *on $\overline{\mathscr{D}_T}$ ($\mathscr{D}_T$) if and only if we have, for every $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T})$ and $\kappa \in (0,1)$,*

$$\max\left\{G(D_X\phi);\ \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t, X, v, D_X\phi), \mathscr{I}_\kappa^\pi(t, X, \phi)\big)\right\} \tag{6.6}$$

whenever $(t,X) \in [0,T) \times \overline{\mathscr{D}}$ $\big((t,X) \in \mathscr{D}_T\big)$ is a global maximum (minimum) relative to $[0,T) \times \overline{\mathscr{D}}$ ($\mathscr{D}_T$) of $v - \phi$.

Proof. We will only prove the statement for subsolutions, as the proof for supersolutions can be proved similarly. Suppose $v \in C_1(\overline{\mathscr{D}_T})$ satisfies (6.6) for all $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t,X) \in \overline{\mathscr{D}_T}$ such that (t,X) is a global maximum relative to $[0,T) \times \overline{\mathscr{D}}$ of $v - \phi$. If (t,X) is a global maximum, $v(s,Y) - v(t,X) \leq \phi(s,Y) - \phi(t,X)$ for all $(s,Y) \in [0,T) \times \overline{\mathscr{D}}$, so $\mathscr{I}^{\pi,\kappa}(t,X,\phi,D_X\phi) \geq \mathscr{I}^{\pi,\kappa}(t,X,v,D_X\phi)$. It follows that

$$\begin{aligned}
F\big(t, X, &D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t,X,\phi)\big)\\
&= F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t,X,\phi,D_X\phi), \mathscr{I}_\kappa^\pi(t,X,\phi)\big)\\
&\geq F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t,X,v,D_X\phi), \mathscr{I}_\kappa^\pi(t,X,\phi)\big),
\end{aligned}$$

and since (6.6) holds, we see that the subsolution inequality (6.2) holds. Therefore v is a viscosity subsolution.

Now we will show the other implication. Let $v \in C_1(\overline{\mathscr{D}_T})$ be a subsolution of (5.11) in $\overline{\mathscr{D}_T}$, and assume (t,X) is a global maximum of $v - \phi$ for some $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T})$. Let $\chi_n : \overline{\mathscr{D}_T} \to [0,1]$ be smooth and satisfy $\chi = 1$ in $\mathscr{N}\big((t,X), x(e^\kappa - 1 - \frac{1}{n})\big) \cap \overline{\mathscr{D}_T}$ and $\chi = 0$ in $\overline{\mathscr{D}_T} \backslash \mathscr{N}\big((t,X), x(e^\kappa - 1)\big)$.

Let $v_n \in C^\infty(\overline{\mathscr{D}_T})$, $n \in \mathbb{N}$, be such that $v_n(s,Y) \leq v(s,Y)$ for all $(s,Y) \in \overline{\mathscr{D}_T}$, and $v_n(s,Y) \to v(s,Y)$ for all $(s,Y) \in \overline{\mathscr{D}_T} \backslash \mathscr{N}\big((t,X), x(e^\kappa - 1)\big)$. Define

$$\psi_n \in C^{1,2,1}\big(\overline{\mathscr{D}_T} \cap C_1(\overline{\mathscr{D}_T})\big)$$

by

$$\psi_n(s,Y) := \chi_n(s,Y)\phi(s,Y) + \big(1 - \chi_n(s,Y)\big)v_n(s,Y)$$

for all $(s,Y) \in \overline{\mathscr{D}_T}$ and $n \in \mathbb{N}$. Observe that $\psi_n = \phi$ in $\mathscr{N}\left((t,X), x(e^\kappa - 1 - \frac{1}{n})\right) \cap \overline{\mathscr{D}_T}$, $\psi_n \to \phi$ in $\mathscr{N}\left((t,X), x\left(e^\kappa - 1\right)\right) \cap \overline{\mathscr{D}_T}\big)$, $\psi_n = v_n$ in $\overline{\mathscr{D}_T} \backslash \mathscr{N}\left((t,X), x(e^\kappa - 1)\right)$, and that (t,X) is a global maximum of $v - \psi_n$. We see from the dominated convergence theorem (Theorem 2.8) that

$$\mathscr{I}^{\pi,\kappa}(t,X,\psi_n,D_X\psi_n) = \mathscr{I}^{\pi,\kappa}(t,X,v_n,D_X\phi) \to \mathscr{I}^{\pi,\kappa}(t,X,v,D_X\phi)$$

and

$$\mathscr{I}_\kappa^\pi(t,X,\psi_n) \to \mathscr{I}_\kappa^\pi(t,X,\phi),$$

so

$$\begin{aligned}
F\big(t, X, &D_X\psi_n, D_X^2\psi_n, \mathscr{I}^\pi(t,X,\psi_n)\big)\\
&= F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t,X,\psi_n,D_X\psi_n), \mathscr{I}_\kappa^\pi(t,X,\psi_n)\big)\\
&\to F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t,X,v,D_X\phi), \mathscr{I}_\kappa^\pi(t,X,\phi)\big).
\end{aligned}$$

It follows that (6.6) holds if v is a subsolution of (5.11) in $\overline{\mathscr{D}_T}$. $\square$

Note that the formulation in Theorem 6.4 and Definition 6.2 are not entirely equivalent, since we assume $v \in C_1(\overline{\mathscr{D}_T})$ in Theorem 6.4. However, we will mostly apply viscosity solution theory to functions satisfying a similar growth condition as V (see Theorem 4.6), and the case $v \notin C_1(\overline{\mathscr{D}_T})$ is therefore not very interesting. In fact, if $v \notin C_1(\overline{\mathscr{D}_T})$ is positive, $v - \phi$ has no global maxima on $\overline{\mathscr{D}_T}$ for any $\phi \in C_1(\overline{\mathscr{D}_T})$, so saying that v is a subsolution does not give any interesting information about v in this case.

The next two lemmas show that we may assume the function ϕ in Definition 6.2 satisfies some additional properties. Note that the limit of ϕ as $x, y \to \infty$ is different in the case of supersolutions and subsolutions. This corresponds to the idea that ϕ should lie above the graph of V for subsolutions, and that ϕ should lie below the graph of V for supersolutions, provided $V - \phi = 0$ at the point where the maximum or minimum, respectively, is taken. The lemmas will be used to prove convergence of the penalty approximations in Chapter 10.

Lemma 6.5. *The function $v : \overline{\mathscr{D}_T} \to [0, \infty)$, $v \in C(\overline{\mathscr{D}_T})$, is a viscosity supersolution of (5.11) on $\mathscr{O}_T \subset \overline{\mathscr{D}_T}$ if and only if*

$$\max\left\{G(D_X\phi);\, \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)\right\} \leq 0 \tag{6.7}$$

for all $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t, X) \in \overline{\mathscr{D}_T}$ that satisfy these conditions:

(1) (t, X) is a global minimum of $v - \phi$ relative to $\mathscr{O}_T$, and there is an $\epsilon' > 0$ such that $(v - \phi)(t', X') > (v - \phi)(t, X) + \epsilon'$ for all other minima (t', X').
(2) $(v - \phi)(t, X) = -a(t, X)$ for some given function $a : \overline{\mathscr{D}_T} \to [0, \infty)$.
(3) ϕ has compact support.

Proof. Let $v : \overline{\mathscr{D}_T} \to [0, \infty)$ be a continuous function. If v is a viscosity supersolution, (6.7) holds for all ϕ satisfying (1)-(3) by Definition 6.2, so we only need to show the opposite implication.

Assume (6.7) holds for all functions ϕ satisfying (1)-(3). We want to show that (6.7) must hold for all ϕ satisfying only

(0) (t, X) is a global minimum of $v - \phi$ relative to $\mathscr{O}_T$.

In Part 1 of the proof we show that (6.7) holds for all functions satisfying (1)-(2), in Part 2 we show that (6.7) holds for all functions satisfying (1), and in Part 3 we show that (6.7) holds for all functions satisfying only (0).

Part 1: Assume ϕ satisfies (1)-(2), and define

$$\phi_\epsilon(t', X') := \phi(t', X')\eta_\epsilon$$

for all $\epsilon > 0$, where $\eta_\epsilon : \overline{\mathscr{D}_T} \to [0, 1]$ satisfies

1. $\eta_\epsilon \in C^\infty(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$,
2. $\eta_\epsilon = 1$ on $\overline{\mathscr{P}_{T,\epsilon}}$, where

$$\mathscr{P}_{T,\epsilon} := \{(t', x', y') \in \mathscr{D}_T : t' \in [0, T),\, x' < x + 1/\epsilon,\, y' < y + 1/\epsilon\},$$

and

3. $\eta(t', x', y') = 0$ for $x' > x + 2/\epsilon$ and $y' > y + 2/\epsilon$.

We see that $\phi_\epsilon \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ must satisfy (2)-(3), and the following argument by contradiction shows that ϕ_ϵ also satisfies (1): Assume

$$(v - \phi_\epsilon)(t', X') \leq (v - \phi_\epsilon)(t, X) + \epsilon'$$

for some $(t', X') \in \overline{\mathscr{D}_T}$. We see immediately that $(t', X') \in \overline{\mathscr{D}_T} \backslash \overline{\mathscr{P}_T}$, because $v - \phi_\epsilon = v - \phi$ on $\overline{\mathscr{P}_T}$. Since ϕ satisfies (1),

$$(v - \phi)(t', X') < (v - \phi)(t, X) + \epsilon'.$$

We also know that

$$(v - \phi_\epsilon)(t', X') < (v - \phi_\epsilon)(t, X) = (v - \phi)(t, X) = -a(t, X) < 0,$$

and since $v(t', X') \geq 0$, this implies that $\phi(t', X') > 0$. We see that

$$\begin{aligned}(v - \phi_\epsilon)(t', X') &= v(t', X') - \phi(t', X')\eta_\epsilon(t', X') \\ &\geq v(t', X') - \phi(t', X') \\ &> (v - \phi)(t, X) + \epsilon' \\ &= (v - \phi_\epsilon)(t, X) + \epsilon',\end{aligned}$$

This is a contradiction to the assumption made above, and it follows that ϕ_ϵ satisfies (1).

We have proved that ϕ_ϵ satisfies (1)-(3), and it follows from the initial assumption that ϕ_ϵ satisfies (6.7). By the dominated convergence theorem, we see that

$$\lim_{\epsilon \to 0} \mathscr{I}^\pi(t, X, \phi_\epsilon) = \mathscr{I}^\pi(t, X, \phi)$$

for all $\pi \in [0, 1]$. This implies that

$$\begin{aligned}&(\phi_\epsilon)_t + F\big(t, X, D_X\phi_\epsilon, D_X^2\phi_\epsilon, \mathscr{I}^\pi(t, X, \phi_\epsilon)\big) \\ &\qquad \to \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)\end{aligned}$$

and

$$G(D_X\phi_\epsilon) \to G(D_X\phi)$$

as $\epsilon \to 0$. It follows that ϕ must satisfy (6.7).

Part 2: Assume ϕ satisfies (1), and define $\phi' \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ by

$$\phi'(t', X') = \phi(t', X') - a(t, X) + v(t, X) - \phi(t, X).$$

We see that ϕ' satisfies (1), since ϕ' and ϕ only differ by a constant, and we also see by insertion that ϕ' satisfies (2). Inequality (6.7) holds for ϕ' by the result in Part 1. We have

$$\phi'_t + F\big(t, X, D_X\phi', D_X^2\phi', \mathscr{I}^\pi(t, X, \phi')\big) = \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)$$

and

$$G(D_X\phi') = G(D_X\phi),$$

and therefore ϕ satisfies (6.7).

Part 3: Now assume ϕ satisfies (0), and define

$$\phi_\epsilon(t', X') := \phi(t', X') + \epsilon\eta_\epsilon(t' - t, X' - X)$$

for all $\epsilon > 0$, where

$$\eta_\epsilon(t', X') := \eta\left(\frac{t'}{\epsilon}, \frac{X'}{\epsilon}\right)$$

and

$$\eta(t', X') := \begin{cases} e^{\frac{1}{x'^2+y'^2+t'^2-1}} & \text{if } |(t', x', y')| < 1, \\ 0 & \text{if } |(t', x', y')| \geq 1. \end{cases}$$

We see that ϕ_ϵ satisfies (1), since η has a strict maximum at $(0, 0, 0)$. By the result in Part 2, ϕ_ϵ satisfies (6.7). We see by direct computations that $\epsilon\eta_\epsilon$, $\epsilon D\eta_\epsilon$ and $\epsilon D^2\eta_\epsilon$ converge towards 0, so by Lemma 6.3, we have

$$\phi'_t + F\big(t, X, D_X\phi', D_X^2\phi', \mathscr{I}^\pi(t, X, \phi')\big) \to \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)$$

and

$$G(D_X\phi') \to G(D_X\phi).$$

It follows that ϕ must satisfy (6.7), and the theorem is proved. □

Theorem 6.6. *The function $v : \overline{\mathscr{D}_T} \to [0, \infty)$, $v \in C(\overline{\mathscr{D}_T})$, is a viscosity subsolution of (5.11) on $\mathscr{O}_T \subset \overline{\mathscr{D}_T}$ if and only if*

$$\max\left\{G(D_X\phi);\ \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)\right\} \geq 0 \tag{6.8}$$

for all $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t, X) \in \overline{\mathscr{D}_T}$ that satisfies these conditions:

(1) (t, X) is a global maximum of $v - \phi$ relative to $\mathscr{O}_T$, and there is an $\epsilon' > 0$ such that $(v - \phi)(t', X') < (v - \phi)(t, X) - \epsilon'$ for all other maxima (t', X').
(2) $(v - \phi)(t, X) = a(t, X)$ for some given function $a : \overline{\mathscr{D}_T} \to [0, \infty)$.
(3) $\lim_{x\to\infty, y\to\infty} \phi(t, x, y)/(x + y) \neq 0$.

Proof. If v is a viscosity subsolution, (6.8) holds for all ϕ satisfying (1)-(3) by Definition 6.2, so we only need to show the opposite implication.

Assume (6.8) holds for all functions ϕ satisfying (1)-(3). We want to show that (6.8) must hold for all ϕ satisfying only

(0) (t, X) is a global maximum of $v - \phi$ relative to $\mathscr{O}_T$.

In Part 1 of the proof we show that (6.8) holds for all functions satisfying (1)-(2), in Part 2 we show that (6.8) holds for all functions satisfying (1), and in Part 3 we show that (6.8) holds for all functions satisfying only (0).

Part 1: Assume ϕ satisfies (1)-(2) and define

$$\phi_\epsilon(t', X') := \phi(t', X') + \epsilon\eta(t', X')$$

for all $\epsilon > 0$, where $\eta : \overline{\mathscr{D}_T} \to [0, 1]$ satisfies the following conditions:

1. $\eta \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$.
2. $\eta(t', X') = 0$ for $x' < x + 1$ and $y' < y + 1$.
3. $\eta(t', X')$ is increasing in x' and y'.

4. $\lim_{x'\to\infty}\eta(t',X')/(x'+y')\neq 0$ and $\lim_{y'\to\infty}\eta(t',X')/(x'+y')\neq 0$ for all $x',y'>0$.

We see immediately that $\phi_\epsilon\in C^{1,2,1}(\overline{\mathscr{D}_T})\cap C_1(\overline{\mathscr{D}_T})$ satisfies (2)-(3), and the following argument by contradiction shows that ϕ_ϵ also satisfies (1): Assume

$$(v-\phi_\epsilon)(t',X')\geq(v-\phi_\epsilon)(t,X)-\epsilon'$$

for some $(t',X')\in\overline{\mathscr{O}_T}$. We see immediately that $x'>x+1$ and $y'>y+1$, because $(v-\phi_\epsilon)(t',X')=(v-\phi)(t,X)$ for $x'<x+1$ and $y'<y+1$. Assume from now on that $x'>x+1$ or $y'>y+1$ Since ϕ satisfies (1)

$$(v-\phi)(t',X')<(v-\phi)(t,X)-\epsilon'.$$

We see that

$$\begin{aligned}(v-\phi_\epsilon)(t',X')&=v(t',X')-\phi(t',X')-\epsilon\eta(t',X')\\&\leq v(t',X')-\phi(t',X')\\&<(v-\phi)(t,X)-\epsilon'\\&=(v-\phi_\epsilon)(t,X)-\epsilon',\end{aligned}$$

This is a contradiction, and it follows that ϕ_ϵ satisfies (1).

We have proved that ϕ_ϵ satisfies (1)-(3), and it follows from the initial assumption that ϕ_ϵ satisfies (6.8).

$$\lim_{\epsilon\to 0}\mathscr{I}^\pi(t,X,\phi_\epsilon)=\lim_{\epsilon\to 0}\mathscr{I}^\pi(t,X,\phi)+\epsilon\mathscr{I}^\pi(t,X,\eta)=\mathscr{I}^\pi(t,X,\phi)$$

for all $\pi\in[0,1]$. This implies that

$$\phi'_t+F\big(t,X,D_X\phi',D_X^2\phi',\mathscr{I}^\pi(t,X,\phi')\big)\to\phi_t+F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)$$

and

$$G(D_X\phi')\to G(D_X\phi)$$

as $\epsilon\to 0$. It follows that ϕ must satisfy (6.8).

Part 2: Assume ϕ satisfies (1), and define $\phi'\in C^{1,2,1}(\overline{\mathscr{D}_T})\cap C_1(\overline{\mathscr{D}_T})$ by

$$\phi'(t',X')=\phi(t',X')-a(t,X)+v(t,X)-\phi(t,X).$$

We see immediately that ϕ' satisfies (1), since ϕ' and ϕ only differ by a constant. We see by insertion that ϕ' also satisfies (2). Inequality (6.8) holds for ϕ' by the result in Part 1. We have

$$\phi'_t+F\big(t,X,D_X\phi',D_X^2\phi',\mathscr{I}^\pi(t,X,\phi')\big)=\phi_t+F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)$$

and

$$G(D_X\phi')=G(D_X\phi),$$

and therefore ϕ satisfies (6.8).

Part 3: Now assume ϕ satisfies (0), and define

$$\phi_\epsilon(t',X')=\phi(t',X')-\epsilon\eta_\epsilon(t'-t,X'-X)$$

for all $\epsilon>0$, where

$$\eta_\epsilon(t',X') := \eta\left(\frac{t'}{\epsilon},\frac{X'}{\epsilon}\right)$$

and

$$\eta(t',X') := \begin{cases} e^{\frac{1}{x'^2+y'^2+t'^2-1}} & \text{if } |(t',x',y')| < 1, \\ 0 & \text{if } |(t',x',y')| \geq 1. \end{cases}$$

We see immediately that ϕ_ϵ satisfies (1), since η has a strict maximum at (t,X). By the result in Part 2, ϕ_ϵ satisfies (6.8). By direct calculations, we see that $\epsilon\eta_\epsilon$, $\epsilon D\eta_\epsilon$ and $\epsilon D^2\eta_\epsilon$ converge towards 0 when $\epsilon \to 0$, so by Lemma 6.3 we see that

$$(\phi_\epsilon)_t + F\big(t,X,D_X\phi_\epsilon,D_X^2\phi_\epsilon,\mathscr{I}^\pi(t,X,\phi_\epsilon)\big) \to (\phi)_t + F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)$$

and

$$G(D_X\phi_\epsilon) \to G(D_X\phi).$$

It follows that ϕ must satisfy (6.8), and the theorem is proved. □

We will now prove two lemmas connecting the theory of viscosity solutions and classical solutions. Lemma 6.7 says that a smooth classical solution of (5.11) also is a viscosity solution, while Lemma 6.8 says that a sufficiently smooth viscosity solution of (5.11) also is a classical solution. Lemma 6.7 will be used to prove that the explicit solution formulas discussed in Chapter 7 are the correct value functions, while Lemma 6.8 only is included for completeness.

It is relatively easy to show that a classical solution of (5.11) on $\mathscr{D}_T$ is a viscosity solution on $\mathscr{D}_T$; the proof is based on studying the sign of the derivative and second derivative of a smooth function at its extremal points.

However, we also wish to show that a classical subsolution on $[0,T)\times\overline{\mathscr{D}}$ is a viscosity subsolution on the whole domain $[0,T)\times\overline{\mathscr{D}}$. The proof given here is inspired by the first remark in Section 3 of [2], where the result is proved for the boundary $y=0$ of a non-singular, time-independent control problem with no jumps in the stochastic process. If $v-\phi$ has a maximum at $(t,X)\in[0,T)\times\overline{\mathscr{D}}\backslash\mathscr{D}_T$, we perturb ϕ to $\phi_\epsilon=\phi+\epsilon\eta_\epsilon$, such that $v-\phi_\epsilon$ has a maximum at the interior point $(t_\epsilon,X_\epsilon)\in\mathscr{D}_T$ and $(t_\epsilon,X_\epsilon)\to(t,X)$ when $\epsilon\to 0$. Using convergence arguments for F and G, we see that ϕ satisfies the viscosity subsolution inequality at (t,X). We have generalized the proof in [2] to include both boundaries $x=0$ and $y=0$, and also take care of the gradient constraint operator G. We also have to be careful considering the convergence of $\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi_\epsilon)$ to $\mathscr{I}^\pi(t,X,\phi)$.

Note that the proof only is valid if (6.9) holds. This corresponds well with results in other parts of the thesis, as we have seen that

$$V_t + F\big(t,X,D_XV,D_X^2V,\mathscr{I}^\pi(t,X,V)\big) \geq 0$$

is expected to hold in a viscosity sense at the boundary $x=0$, while

$$G(D_XV) \geq 0$$

is expected to hold in a viscosity sense at the boundary $y=0$. See for example the discussion after Definition 6.2, the explicit solution formula derived in Section 7.1, and the discussion of boundary conditions in Section 11.3.

Lemma 6.7. *If $v \in C^{1,2,1}(\overline{\mathscr{D}_T})$ is a classical solution of the subsolution (supersolution) inequality* (6.1) *in* $[0,T)\times\overline{\mathscr{D}}$ *($\mathscr{D}_T$), it is a viscosity subsolution (supersolution) of* (5.11) *in* $[0,T)\times\overline{\mathscr{D}}$ *($\mathscr{D}_T$). In the subsolution case we also need the additional assumption*

$$v_t + F\big(t,0,y,D_Xv,D_X^2v,\mathscr{I}^\pi(t,0,y,v)\big) \geq 0 \tag{6.9}$$

for all $y \in [0,\infty)$.

Proof. We will first prove that v is a viscosity supersolution of (5.11) in $\mathscr{D}_T$ if it is a classical solution of the supersolution inequality (6.1). Suppose $(t,X) \in \mathscr{D}_T$ is a global minimum of $v-\phi$. Since (t,X) is a minimum, we have $D_X\phi(t,X) = D_Xv(t,X)$, $\phi_t(t,X) \leq v_t(t,X)$, $\phi_{xx}(t,X) \leq v_{xx}(t,X)$ and $\mathscr{I}^\pi(t,X,\phi) \leq \mathscr{I}^\pi(t,X,v)$. The reason for the inequality sign in $\phi_t(t,X) \leq v_t(t,X)$ is that t might be 0, and we do not get an inequality sign in $D_X\phi(t,X) = D_Xv(t,X)$, because X is an interior point of $\mathscr{D}$. We see that

$$G(D_X\phi) = G(D_Xv)$$

and

$$\begin{aligned}\phi_t+&F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\\ &\leq v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big).\end{aligned}$$

Using that v is a classical solution of (5.11), we get

$$\begin{aligned}\max&\big\{G(D_X\phi);\ \phi_t + F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\big\}\\ &\leq \max\big\{G(D_Xv);\ v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big)\big\}\\ &\leq 0,\end{aligned}$$

so v is a supersolution of (5.11) in $\mathscr{D}_T$.

Now we will prove that v is a subsolution of (5.11) in $[0,T)\times\overline{\mathscr{D}}$. We start by proving that v is a subsolution of (5.11) in $\mathscr{D}_T$. Suppose $(t,X) \in \mathscr{D}_T$ is a global maximum of $v-\phi$. Since (t,X) is a global maximum we have $D_X\phi(t,X) = D_Xv(t,X)$, $\phi_t(t,X) \geq v_t(t,X)$, $\phi_{xx}(t,X) \geq v_{xx}(t,X)$ and $\mathscr{I}^\pi(t,X,\phi) \geq \mathscr{I}^\pi(t,X,v)$. We see that

$$G(D_X\phi) = G(D_Xv)$$

and

$$\begin{aligned}\phi_t+&F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\\ &\geq v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big).\end{aligned}$$

We have

$$\max\big\{G(D_Xv);\ v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big)\big\} \geq 0,$$

since v is a classical solution of the subsolution inequality (6.1). Therefore v is also a viscosity subsolution of (5.11) on $\mathscr{D}_T$:

$$\begin{aligned}\max&\big\{G(D_X\phi);\ \phi_t + F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\big\}\\ &\geq \max\big\{G(D_Xv);\ v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^\pi(t,X,v)\big)\big\}\\ &\geq 0.\end{aligned}$$

We will now prove that v is a subsolution on the whole domain $[0,T)\times\overline{\mathscr{D}}$. Suppose $(t,X)\in[0,T)\times\overline{\mathscr{D}}$ is a strict global maximum of $v-\phi$, and define

$$\phi_\epsilon(t',X')=\phi(t',X')-\epsilon(\ln x'+\ln y')$$

for $\epsilon>0$. For all $\epsilon>0$ $v-\phi_\epsilon$ has a local maximum $(t_\epsilon,X_\epsilon)\in\mathscr{D}_T$, and $(t_\epsilon,X_\epsilon)\to(t,X)$ as $\epsilon\to0$. ϕ_ϵ is not $C^{1,2,1}(\overline{\mathscr{D}_T})\cap C_1(\overline{\mathscr{D}_T})$, but we can modify it outside a neighbourhood of (t_ϵ,X_ϵ) such that it becomes $C^{1,2,1}(\overline{\mathscr{D}_T})\cap C_1(\overline{\mathscr{D}_T})$. We can write ϕ_ϵ on the form

$$\phi_\epsilon=\phi+\epsilon\eta_\epsilon,$$

where $\eta_\epsilon\in C^{1,2,1}(\overline{\mathscr{D}_T})\cap C_1(\overline{\mathscr{D}_T})$ and $\eta(t',X')=\ln x'+\ln y'$ in a neighbourhood around (t_ϵ,X_ϵ) and everywhere else except close to the boundary of $\overline{\mathscr{D}}$. We can assume

(1) $\eta(t',X')=\ln x'+\ln y'$ except when $x'<x_\epsilon+x_\epsilon(e^{-1}-1)$ or $y'<y_\epsilon+y_\epsilon(e^{-1}-1)$,
(2) $\eta(t',X')\geq\ln x'+\ln y'$ for all $(t',X')\in\overline{\mathscr{D}_T}$, and
(3) $\eta_{xx}(t',X')\leq0$ for all $(t',X')\in\overline{\mathscr{D}_T}$.

By the result in the previous paragraph, we know that

$$\max\left\{G\big(D_X\phi_\epsilon(t_\epsilon,X_\epsilon)\big);\,(\phi_\epsilon)_t+F\big(t_\epsilon,X_\epsilon,D_X\phi_\epsilon,D_X^2\phi_\epsilon,\mathscr{I}^\pi\big(t_\epsilon,X_\epsilon,\phi_\epsilon\big)\big)\right\}\geq0 \quad (6.10)$$

for all $\epsilon>0$.

In the case $x\neq0$, we have

$$G(D_X\phi)\geq G(D_Xv), \quad (6.11)$$

since $\phi_x=v_x$ and $v_y\leq\phi_y$.

Now we want to prove in the general case, i.e., both x and y can be 0, that

$$F\big(t_\epsilon,X_\epsilon,D_X\phi_\epsilon,D_X^2\phi_\epsilon,\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi_\epsilon)\big)\to F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big). \quad (6.12)$$

By Lemma 6.3,

$$F\big(t_\epsilon,X_\epsilon,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi)\big)\to F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big),$$

so what remains to be proved is that

$$|F(t_\epsilon,X_\epsilon,D_X\phi_\epsilon,D_X^2\phi_\epsilon,\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi_\epsilon)-F(t_\epsilon,X_\epsilon,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi)|\to0$$

as $\epsilon\to0$. We have $x_\epsilon(\phi_\epsilon)_x(t_\epsilon,X_\epsilon)=x_\epsilon\phi_x(t_\epsilon,X_\epsilon)-\epsilon$ and $y_\epsilon(\phi_\epsilon)_y(t_\epsilon,Y_\epsilon)=y_\epsilon\phi_y(t_\epsilon,Y_\epsilon)-\epsilon$, so it is sufficient to prove that

$$|\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi_\epsilon)-\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi)|\to0.$$

Note that this does not follow from Lemma 6.3, since ϕ_ϵ does not converge to ϕ on the whole domain $\overline{\mathscr{D}_T}$. We have

$$\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi_\epsilon)-\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\phi)=-\epsilon\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\eta_\epsilon),$$

so it is sufficient to prove that the $\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\eta_\epsilon)$ is bounded as $\epsilon\to0$. To prove that $\mathscr{I}^\pi(t_\epsilon,X_\epsilon,\eta_\epsilon)$ is bounded, we consider the cases $|z|<1$, $z>1$ and $z<-1$ separately. For $|z|<1$

$$\eta_\epsilon(t_\epsilon, x_\epsilon + x_\epsilon\pi(e^z-1), y_\epsilon) = \ln\big(x_\epsilon + x_\epsilon\pi(e^z-1)\big) + \ln y_\epsilon$$

by assumption (1), so

$$\begin{aligned}\int_{|z|<1} &\eta_\epsilon(t_\epsilon, x_\epsilon + \pi(e^z-1)x_\epsilon, y_\epsilon) - \eta_\epsilon(t_\epsilon, X_\epsilon) - \pi(e^z-1)(\eta_\epsilon)_x(t_\epsilon, X_\epsilon)\,\nu(dz)\\ &= \int_{|z|<1} \ln\big(1+\pi(e^z-1)\big) - \pi(e^z-1)\,\nu(dz),\end{aligned}$$

which is finite by (3.1), as the integrand is of order $O(z^2)$.

Now we consider $z > 1$. By using (1), we get

$$\begin{aligned}\int_{z>1} &\eta_\epsilon(t_\epsilon, x_\epsilon + \pi x_\epsilon(e^z-1)x_\epsilon, y_\epsilon) - \eta_\epsilon(t_\epsilon, X_\epsilon) - \pi(e^z-1)(\eta_\epsilon)_x(t_\epsilon, X_\epsilon)\,\nu(dz)\\ &= \int_{z>1} \ln\big(1+\pi(e^z-1)\big) + \pi(e^z-1)\,\nu(dz),\end{aligned}$$

which is finite by (3.1), as the integrand is of order e^z.

Now we consider the case $z < -1$. By doing a Taylor expansion around $(t_\epsilon, x_\epsilon, y_\epsilon)$ and using (3), we get

$$\begin{aligned}\eta_\epsilon&(t_\epsilon, x_\epsilon + \pi x_\epsilon(e^z-1), y_\epsilon) - \eta_\epsilon(t_\epsilon, X_\epsilon) - \pi x_\epsilon(e^z-1)(\eta_\epsilon)_x(t_\epsilon, X_\epsilon)\\ &= \frac{1}{2}\pi^2 x_\epsilon^2(e^z-1)^2(\eta_\epsilon)_{xx}(t_\epsilon, a, y_\epsilon)\\ &\le 0,\end{aligned}$$

where a is some number between x_ϵ and $x_\epsilon + \pi(e^z-1)x_\epsilon$. By (2), we get

$$\begin{aligned}\eta_\epsilon&(t_\epsilon, x_\epsilon + \pi(e^z-1)x_\epsilon, y_\epsilon) - \eta_\epsilon(t_\epsilon, X_\epsilon) - \pi(e^z-1)(\eta_\epsilon)_x(t_\epsilon, X_\epsilon)\\ &\ge \ln\big(1+\pi(e^z-1)\big) + \pi(e^z-1).\end{aligned}$$

If $\pi = 1$,

$$\ln\big(1+\pi(e^z-1)\big) + \pi(e^z-1) = O(z)$$

is not necessarily integrable with respect to ν on $(-\infty, -1]$. If $\pi \neq 1$, on the other hand,

$$\ln\big(1+\pi(e^z-1)\big) + \pi(e^z-1) = O(e^z-1)$$

is integrable on $(-\infty, -1]$ with respect to ν. However we know that $\pi \nrightarrow 1$ as $\epsilon \to 0$, because $\pi \in [0,1]$ always is chosen to maximize an expression involving

$$\int_{z<-1} \eta_\epsilon(t_\epsilon, x_\epsilon + \pi(e^z-1)x_\epsilon, y_\epsilon) - \eta_\epsilon(t_\epsilon, X_\epsilon) - \pi(e^z-1)(\eta_\epsilon)_x(t_\epsilon, X_\epsilon)\,\nu(dz). \tag{6.13}$$

By (3.1) and the dominated convergence theorem 2.8, we see that the integral (6.13) is bounded as $\epsilon \to 0$, and we have proved (6.12).

Equations (6.9), (6.10), (6.11), (6.12) and $\phi_t \ge v_t$ imply that

$$\max\big\{G(D_X\phi);\ \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)\big\} \ge 0.$$

$\square$

The lemma below proves the opposite implication of the lemma above. We do not have to treat the boundary of $\overline{\mathscr{D}_T}$ as carefully in this proof, as we can use continuity arguments to show directly that the subsolution/supersolution inequality holds at the boundary. However, we need to use the formulation of viscosity solutions given in Theorem 6.4 to prove the result for interior points of $\mathscr{D}_T$.

Lemma 6.8. *Let $v \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ be a viscosity subsolution (supersolution) of* (5.11) *on $(0,T) \times \mathscr{D}_T$. Then v is a classical subsolution (supersolution) of* (5.11) *on $\overline{\mathscr{D}_T}$.*

Proof. The result will only be proved for subsolutions, as the proof for supersolutions is very similar. Let $(t,X) \in (0,T) \times \mathscr{D}$, and choose a function $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ such that $v - \phi$ has a global maximum at (t,X). Such a function exists since $v \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$. Since v is a viscosity subsolution, we have

$$\max\left\{G(D_X\phi);\ \phi_t + F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^{\pi,\kappa}(t,X,v,D_X\phi),\mathscr{I}_\kappa^{\pi}(t,X,\phi)\big)\right\} \geq 0.$$

for all $\kappa > 0$ by Theorem 6.4. We see that $D_Xv = D_X\phi$, $D_X^2 v = D_X^2\phi$ and $v_t = \phi_t$, since (t,X) is in the interior of $\mathscr{D}_T$. It follows that

$$G(D_Xv) = G(D_X\phi)$$

and

$$\begin{aligned}\phi_t + F\big(t,&X,D_X\phi,D_X^2\phi,\mathscr{I}^{\pi,\kappa}(t,X,v,D_X\phi),\mathscr{I}_\kappa^{\pi}(t,X,\phi)\big)\\ &\to v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^{\pi,\kappa}(t,X,v,D_Xv),\mathscr{I}_\kappa^{\pi}(t,X,v)\big)\\ &= v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^{\pi}(t,X,v)\big)\end{aligned}$$

as $\kappa \to 0$, so

$$\max\left\{G(D_Xv);\ v_t + F\big(t,X,D_Xv,D_X^2v,\mathscr{I}^{\pi}(t,X,v)\big)\right\} \geq 0. \tag{6.14}$$

By continuity, we see that (6.14) holds for all $(t,X) \in \overline{\mathscr{D}_T}$, so v is a classical subsolution of (5.11) on $\overline{\mathscr{D}_T}$. $\square$

6.2 Results needed for comparison principle

The proof of uniqueness of viscosity solutions of (5.11) and (4.2) is based on a comparison principle that will be proved in Theorem 6.15. For HJB equations of first order, the formulation given in Theorem 6.4 is an appropriate definition of viscosity solutions for proving the comparison principle. However, due to the Brownian motion component of the Lévy process, the HJB equation (5.11) is of second order. To deal with the second order term of (5.11), we need another characterization of viscosity solutions.

In the local case, i.e., the case where there is no integral term in the HJB equation, we can give a characterization of viscosity solutions by using so-called superdifferentials/superjets and subdifferentials/subjets, see Definition 6.11 below. If we let $\tilde{\mathscr{D}} \subseteq \mathbb{R}^n$ and let $\tilde{F}(X,v,D_Xv,D_X^2v) = 0$ be an HJB equation, it turns out that $v \in C(\tilde{\mathscr{D}})$ is a supersolution (subsolution) of the HJB equation if and only if, for all $(P,A) \in J_{\tilde{\mathscr{D}}}^{2,-(+)}v(X)$,

$\tilde{F}(X, v, P, A) \leq 0\,(\geq 0)$. By using the maximum principle stated in Theorem 3.2 in [15], it is possible to show a comparison principle in the second-order case.

In [8] the authors try to generalize this result to the non-local case. However, some of the convergence arguments are not done carefully enough, and the proof of the comparison principle in [8] is therefore not entirely correct, see the end of this section for more information. In [35] a new characterization of viscosity solutions by subjets and superjets is given, see Lemma 6.12 below. However, note that this lemma is not a direct extension of the result in the local case, as the information $(P, A) \in J^{2,-(+)}_{\tilde{\mathscr{D}}} v(t, X)$ alone is not enough to guarantee that

$$\max\left\{G(\widehat{P});\ \tilde{P} + F\big(t, X, \widehat{P}, \widehat{A}, \mathscr{I}^{\pi,\kappa}(t, X, v, \widehat{P}), \mathscr{I}^{\pi}_{\kappa}(t, X, \phi)\big)\right\} \leq 0\,(\geq 0)$$

for some $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$, where $\tilde{P}$ denotes the part of P corresponding to t, $\widehat{P}$ denotes the part of P corresponding to X, and $\widehat{A}$ denotes the part of A corresponding to X.

Another feature of the proof of the comparison principle, is that we work with *strict* supersolutions. Lemma 6.10 shows that we can find strict supersolutions of (5.11) arbitrarily close to any supersolution of (5.11) in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$ for some $\gamma' \in (0, 1)$. First we define what we mean by strict subsolutions and supersolutions.

Definition 6.9 (Strict subsolutions and supersolutions). *Let $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$. Any $v \in C(\overline{\mathscr{D}_T})$ is a strict supersolution (subsolution) of* (5.11) *in $\mathscr{O}_T$ if and only if we have, for every $(t, X) \in \mathscr{O}_T$ and any $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ such that (t, X) is a global minimum of $v - \phi$ relative to $\mathscr{O}_T$,*

$$\max\left\{G(D_X\phi);\ \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi}(t, X, \phi)\big)\right\} \leq -\alpha\,(\geq \alpha) \tag{6.15}$$

for some constant $\alpha > 0$. Repeating the proof of Lemma 6.4, we can equivalently replace the left-hand side of (6.15) *by the left-hand side of* (6.6) *if $v \in C_1(\overline{\mathscr{D}_T})$ and $\mathscr{O}_T = \mathscr{D}_T$ ($\mathscr{O}_T = \overline{\mathscr{D}_T}$).*

Now we will prove a lemma showing the existence of strict supersolutions arbitrarily close to non-strict supersolutions. It corresponds to a similar lemma in the infinite-horizon case in [9]. The function $\hat{w} = (K + \chi^{\bar{\gamma}})$ in that lemma is the same as the function w in the lemma below, except for the factor $e^{-\delta t}$. Multiplying by this factor makes $F\big(t, X, D_X w, D_X^2 w, \mathscr{I}^{\pi}(t, X, w)\big)$ similar to the corresponding function

$$\begin{aligned}\widehat{F}\big(X, \hat{w}, &D_X\hat{w}, D_X^2\hat{w}, \widehat{\mathscr{I}^{\pi}}(X, \hat{w})\big)\\ &= U(y) - \delta\hat{w} - \beta y\hat{w}_y\\ &\quad + \max_{\pi\in[0,1]}\left[\big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big)x\hat{w}_x + \frac{1}{2}(\sigma\pi x)^2\hat{w}_{xx} + \widehat{\mathscr{I}^{\pi}}(X, \hat{w})\right]\end{aligned}$$

in [9], where

$$\widehat{\mathscr{I}^{\pi}}(X, \hat{w}) = \int_{\mathbb{R}\setminus\{0\}} \hat{w}\big(x + x\pi(e^z - 1), y\big) - \hat{w}(x, y) - \pi x(e^z - 1)\hat{w}_x(x, y)\,\nu(dz).$$

Lemma 6.10. *For $\gamma' > 0$ such that $\delta > k(\gamma')$, let $v \in C_{\gamma'}(\overline{\mathscr{D}_T})$ be a supersolution of* (5.11) *in $\mathscr{D}_T$. Choose $\bar{\gamma} > \max\{\gamma, \gamma'\}$ such that $\delta > k(\bar{\gamma})$, and let*

$$w = (K + \chi^{\bar{\gamma}})e^{-\delta t},\ \chi(X) = \left(1 + x + \frac{y}{2\beta}\right).$$

Then for K sufficiently large, $w \in C^\infty(\mathscr{D}) \cap C_{\bar{\gamma}}(\bar{\mathscr{D}}_T)$ is a strict supersolution of (5.11) *in any bounded set $\mathscr{O}_T \subset \mathscr{D}_T$. Moreover, for $\theta \in (0,1]$, the function*

$$v^\theta = (1-\theta)v + \theta w \in C_{\bar{\gamma}}(\overline{\mathscr{D}_T}),$$

is a strict supersolution of (5.11) *in any bounded set $\mathscr{O}_T \subseteq \mathscr{D}_T$.*

Proof. First we want to prove that

$$\max\left\{G(D_X w);\ w_t + F\big(t, X, D_X w, D_X^2 w, \mathscr{I}^\pi(t,X,w)\big)\right\} \leq -f \tag{6.16}$$

for some $f \in C(\overline{\mathscr{D}_T})$ that is strictly positive. This will imply the first part of Lemma 6.10 by Theorem 6.7. To prove (6.16), observe that

$$G(D_X w) = -e^{-\delta t}\frac{\bar{\gamma}}{2}\chi^{\bar{\gamma}-1}. \tag{6.17}$$

Since $\frac{x}{\chi}, \pi\frac{x}{\chi} \in [0,1]$, we also have

$$\begin{aligned}
w_t + &F\big(t, X, D_X w, D_X^2 w, \mathscr{I}^\pi(t,X,w)\big)\\
&= e^{-\delta t}\Bigg[U(y) - \delta(K + \chi^{\bar{\gamma}}) - \frac{1}{2}y\bar{\gamma}\chi^{\bar{\gamma}-1}\\
&\quad + \max_{\pi\in[0,1]}\Bigg[\bar{\gamma}\big(\hat{r} + (\hat{\mu}-\hat{r})\pi\big)x\chi^{\bar{\gamma}-1} + \frac{1}{2}\bar{\gamma}(\bar{\gamma}-1)(\sigma\pi x)^2\chi^{\bar{\gamma}-1}\\
&\quad \int_{\mathbb{R}\setminus\{0\}}\Big(\chi + \pi x(e^z - 1)\Big)^{\bar{\gamma}} - \chi^{\bar{\gamma}} - \bar{\gamma}\pi x\chi^{\bar{\gamma}-1}(e^z-1)\,\nu(dz)\Bigg]\Bigg]\\
&= e^{-\delta t}\Bigg[U(y) - \delta K - \frac{1}{2}y\bar{\gamma}\chi^{\bar{\gamma}-1}\\
&\quad + \Bigg(-\delta + \max_{\pi\in[0,1]}\Bigg[\bar{\gamma}(\hat{r} + (\hat{\mu}-\hat{r})\pi)\frac{x}{\chi} + \frac{1}{2}\bar{\gamma}(\bar{\gamma}-1)\left(\sigma\pi\frac{x}{\chi}\right)^2\\
&\quad \int_{\mathbb{R}\setminus\{0\}}\left(\left(1 + \pi\frac{x}{\chi}(e^z-1)\right)^{\bar{\gamma}} - 1 - \bar{\gamma}\pi\frac{x}{\chi}(e^z-1)\right)\nu(dz)\Bigg]\Bigg)\chi^{\bar{\gamma}}\Bigg]\\
&\leq e^{-\delta t}\Bigg[U(y) - \delta K + \Bigg(-\delta + \max_{\pi\in[0,1]}\Bigg[\bar{\gamma}(\hat{r} + (\hat{\mu}-\hat{r})\pi) + \frac{1}{2}\bar{\gamma}(\bar{\gamma}-1)(\sigma\pi)^2\\
&\quad + \int_{\mathbb{R}\setminus\{0\}}\big(1 + \pi(e^z-1)\big)^{\bar{\gamma}} - 1 - \bar{\gamma}\pi(e^z-1)\,\nu(dz)\Bigg]\Bigg)\chi^{\bar{\gamma}}\Bigg]\\
&\leq e^{-\delta t}\Big(U(y) - \delta K - \big(\delta - k(\bar{\gamma})\big)\chi^{\bar{\gamma}}\Big)\\
&\leq -1.
\end{aligned}$$

The last inequality follows by choosing for example

$$K = \frac{1}{\delta}\left(e^{\delta T} + \sup_{(t,X)\in\overline{\mathscr{D}_T}} \Big[U(y) - (\delta - k(\bar\gamma))\chi^{\bar\gamma}\Big]\right).$$

Note that K is finite, because $\delta > k(\bar\gamma)$ and $\bar\gamma > \gamma$. By defining

$$f(t,X) = \min\left\{1;\ \frac{\bar\gamma}{2}\chi^{\bar\gamma-1}(X)\right\},$$

we see that (6.16) holds.

Now we will prove that v^θ is a strict supersolution of (5.11) for all $\theta \in (0,1]$. Note that for any $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T})$, (t,X) is a global minimum of $v-\phi$ if and only if (t,X) is a global minimum of $v^\theta - \phi^\theta$, where $\phi^\theta = (1-\theta)\phi + \theta w$. Since v is a supersolution of (5.11) in $\mathscr{O}_T$ and G is linear, we have

$$G(D_X\phi^\theta) \le -\theta\frac{\bar\gamma}{2}\chi^{\bar\gamma-1}. \tag{6.18}$$

Let π^* denote the maximizing value of π when evaluating F with ϕ^θ. We have

$$\begin{aligned}
&\phi^\theta_t + F\big(t,X,D_X\phi^\theta,D_X^2\phi^\theta,\mathscr{I}^\pi(t,X,\phi^\theta)\big)\\
&\quad= (1-\theta)U(y)e^{-\delta t} + (1-\theta)\phi_t + (1-\theta)\beta y\phi_y\\
&\qquad+ (1-\theta)(\hat r + (\hat\mu + \hat r)\pi^*)x\phi_x + (1-\theta)\frac{1}{2}(\sigma\pi^* x)^2\phi_{xx}\\
&\qquad+ (1-\theta)\mathscr{I}^{\pi^*}(t,X,\phi) + \theta U(y)e^{-\delta t} + \theta w_t + \theta\beta y w_y\\
&\qquad+ \theta(\hat r + (\hat\mu + \hat r)\pi^*)x w_x + \theta\frac{1}{2}(\sigma\pi^* x)^2 w_{xx} + \theta\mathscr{I}^{\pi^*}(t,X,w)\\
&\quad\le (1-\theta)F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\\
&\qquad+ \theta F\big(t,X,D_Xw,D_X^2w,\mathscr{I}^\pi(t,X,w)\big)\\
&\quad\le -\theta f.
\end{aligned} \tag{6.19}$$

Combining (6.18) and (6.19), we see that

$$\max\left\{G(D_X\phi^\theta);\ \phi^\theta_t + F\big(t,X,D_X\phi^\theta,D_X^2\phi^\theta,\mathscr{I}^\pi(t,X,\phi^\theta)\big)\right\} \le -\theta f.$$

For any bounded subset of $\overline{\mathscr{D}_T}$, there is a constant $\alpha > 0$ such that $\theta f \ge \alpha$. It follows that v^θ is a strict supersolution of (5.11) on any bounded subset of $\overline{\mathscr{D}_T}$. □

Now we will return to the problem of finding a characterization of supersolutions and subsolutions using subjets and superjets. First we define what we mean by superjets and subjets. The definition below is from [15].

Definition 6.11. *Let $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$, $v \in C(\mathscr{O}_T)$ and $(t,X) \in \mathscr{O}_T$. The second order superjet (subjet) $J^{2,+(-)}_{\mathscr{O}_T}v(t,X)$ is the set of $(P,A) \in \mathbb{R}^3 \times \mathbb{S}^3$ such that*

$$\begin{aligned}
v(s,Y) \le (\ge)\, v(t,X) &+ \Big\langle P, (s,Y)-(t,X)\Big\rangle\\
&+ \frac{1}{2}\Big\langle A\big((s,Y)-(t,X)\big), (s,Y)-(t,X)\Big\rangle + o\big(|(t,X)-(s,Y)|^2\big)
\end{aligned}$$

when $(s,Y) \to (t,X)$ for $(s,Y) \in \mathscr{O}_T$, where $\langle\cdot,\cdot\rangle$ denotes inner product. The closure $\bar J^{2,+(-)}_{\mathscr{O}_T}v(t,X)$ is the set of (P,A) for which there exists a sequence $\{(t_n,X_n,P_n,A_n)\}_{n\in\mathbb{N}}$, $(P_n,A_n) \in J^{2,+(-)}_{\mathscr{O}_T}v(t_n,X_n)$, such that

$$\big(t_n, X_n, v(t_n, X_n), P_n, A_n\big) \to \big(t, X, v(t, X), P, A\big)$$

as $n \to \infty$.

The lemma below is a characterization of viscosity solutions using subjets and superjets, and is from [35]. Our formulation of the lemma uses *strict* supersolutions. If we had defined $\overline{v}$ to be a non-strict supersolution, we had obtained the same result, but with $f = 0$. Another difference between the formulation in [35] and here, is that we define $\underline{v}$ to be a subsolution also at the boundary of the domain, and hence allow (t_1^*, X_1^*) to be on the boundary of $\mathscr{D}_T$. The lemma stated below is equivalent to the corresponding lemma in [35], except that the HJB equation is written on a more general form in [35].

Lemma 6.12. *Let* $\underline{v} \in C_2(\overline{\mathscr{D}_T}) \cap USC(\overline{\mathscr{D}_T})$ *be a subsolution of* (5.11) *on* $[0, T) \times \overline{\mathscr{D}}$, *and let* $\overline{v} \in C_2(\overline{\mathscr{D}_T}) \cap LSC(\overline{\mathscr{D}_T})$ *be a strict supersolution of* (5.11) *on* $\mathscr{D}_T$. *Let* f *be a strictly positive function that satisfies*

$$\begin{aligned}-f(t_2, X_2) > \max\big\{&G(D_X\psi);\ \psi_t + F\big(t_2, X_2, D_X\psi, D_X^2\psi,\\ &\mathscr{I}^{\pi,\kappa}(t_2, X_2, \overline{v}, D_X\psi), \mathscr{I}^{\pi}_{\kappa}(t_2, X_2, \psi)\big)\big\}\end{aligned}$$

for all $\psi \in C^{1,2,1}(\overline{D_T}) \cap C_1(\overline{D_T})$ *such that* $\overline{v} - \psi$ *has a global minimum at* $(t_2, X_2) \in \mathscr{D}_T$. *Let* $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T} \times \overline{\mathscr{D}_T})$ *and* $\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big) \in \big([0, T) \times \overline{\mathscr{D}}\big) \times \mathscr{D}_T$ *be such that* $\Phi : \overline{\mathscr{D}_T} \times \overline{\mathscr{D}_T} \to \mathbb{R}$ *has a global maximum at* $\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)$, *where*

$$\Phi\big((t_1, X_1), (t_2, X_2)\big) := \underline{v}(t_1, X_1) - \overline{v}(t_2, X_2) - \phi\big((t_1, X_1), (t_2, X_2)\big).$$

Furthermore, assume that in a neighbourhood of $\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)$, *there are continuous functions* $g_0 : \mathbb{R}^6 \to \mathbb{R}$, $g_1, g_2 : \mathbb{R}^3 \to \mathbb{S}^3$ *with* $g_0\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big) > 0$ *satisfying*

$$D^2\phi \le g_0\big((t_1, X_1), (t_2, X_2)\big) \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \begin{pmatrix} g_1(t_1, X_1) & 0 \\ 0 & g_2(t_2, X_2) \end{pmatrix}.$$

Then, for any $\varsigma \in (0, 1)$ *and* $\kappa > 0$, *there exist two matrices* $A_1, A_2 \in \mathbb{S}^3$ *satisfying*

$$\begin{aligned}-\frac{2g_0\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)}{1-\varsigma}\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \le \begin{pmatrix} A_1 & 0 \\ 0 & -A_2 \end{pmatrix} - \begin{pmatrix} g_1(t_1^*, X_1^*) & 0 \\ 0 & g_2(t_2^*, X_2^*) \end{pmatrix}\\ \le \frac{g_0\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)}{\varsigma}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix},\end{aligned} \tag{6.20}$$

such that

$$\begin{aligned}0 \le \max\big\{&G(D_X\phi);\ \phi_t + F\big(t_1^*, X_1^*, D_X\phi, \widehat{A}_1,\\ &\mathscr{I}^{\pi,\kappa}(t_1^*, X_1^*, \underline{v}, D_X\phi), \mathscr{I}^{\pi,\kappa}(t_1^*, X_1^*, \phi)\big)\big\}\end{aligned} \tag{6.21}$$

and

$$\begin{aligned}-f > \max\big\{&G(-D_X\phi);\ -\phi_t + F\big(t_2^*, X_2^*, -D_X\phi, \widehat{A}_2,\\ &\mathscr{I}^{\pi,\kappa}(t_2^*, X_2^*, \overline{v}, -D_X\phi), \mathscr{I}^{\pi}_{\kappa}(t_2^*, X_2^*, -\phi)\big)\big\},\end{aligned} \tag{6.22}$$

where $\widehat{A}_1$ *and* $\widehat{A}_2$ *are the parts of* A_1 *and* A_2, *respectively, corresponding to the* X *variable.*

In the proof of Lemma 6.12 it is shown that $(D\phi(t_1^*, X_1^*, t_2^*, X_2^*), A_1) \in \bar{J}^{2,+}\underline{v}(t_1^*, X_1^*)$ and $(-D\phi(t_1^*, X_1^*, t_2^*, X_2^*), A_2) \in \bar{J}^{2,+}\overline{v}(t_2^*, X_2^*)$.

We have now stated all the results that will be needed it later chapters, but will end the section with a digression concerning the proof of the comparison principle, and the difference between the approach in [8] and [35]. In [8] the authors consider an infinite-horizon version of our problem, and they show a comparison principle by assuming the existence of functions ϕ, ϕ_n, $n \in \mathbb{N}$, with some wanted properties. The existence of the functions ϕ, ϕ_n is not shown by a careful enough argument, and the proof in [8] is therefore not entirely correct. In [35] an alternative proof of a comparison principle is given. However, it is not proved in [35] that functions ϕ and ϕ_n with the wanted properties do not exist, so [35] leaves the question of their existence open.

It will be shown here that functions ϕ and ϕ_n do, in fact, not exist if we make some additional assumptions on the value function $\widehat{V}$. The assumptions we will make on the value function, are that $\bar{J}^{2,-}\widehat{V}(X)\backslash J^{2,-}\widehat{V}(X)$ is non-empty for some $X \in \mathscr{D}$, and that $\widehat{V}$ is not differentiable at X. The proof will be done by contradiction. We will assume that some functions ϕ and ϕ_n with the wanted properties exist, and will show that this implies that the value function $\widehat{V}$ is differentiable. Since we assumed $\widehat{V}$ is not differentiable, we have obtained a contradiction, and it follows that functions ϕ and ϕ_n not exist. First we will describe the procedure for proving the comparison principle in [9]. Then we will prove a lemma, Lemma 6.13, stating that the approach in [9] is not correct if we make some assumptions on $\widehat{V}$.

It is shown in [9] that the value function $\widehat{V} : \overline{\mathscr{D}} \to \mathbb{R}$ is concave, non-decreasing and continuous. The proof of the comparison principle in [9] is based on the following statement: Given any $(p, A) \in \bar{J}^{2,-}\widehat{V}(X)$ for some $X \in \overline{\mathscr{D}}$, there are functions $\phi, \phi_n \in C^2(\overline{\mathscr{D}})$ and $(p_n, A_n) \in J^{2,-}\widehat{V}(X)$, $n \in \mathbb{N}$, such that

(1) $\phi_n \to \phi$, $D\phi_n \to D\phi$, $D^2\phi_n \to D^2\phi$,
(2) $p_n = D\phi_n(X_n)$ and $A_n = D^2\phi_n(X_n)$,
(3) $(p_n, A_n) \in J^{2,-}\widehat{V}(X_n)$,
(4) $(X_n, p_n, A_n) \to (X, p, A)$, and
(5) X_n is a global minimum of $\widehat{V} - \phi_n$.

Since $(p, A) \in \bar{J}^{2,-}\widehat{V}(X)$, there is a sequence $(X_n, p_n, A_n) \to (X, p, A)$ such that $(p_n, A_n) \in J^{2,-}\widehat{V}(X_n)$, so we can easily find (X_n, p_n, A_n) satisfying (3) and (4). The existence of functions ϕ_n satisfying (2) and (5) can be proved by a similar argument as in Lemma 4.1 in [25]. However, it is not necessarily possible to find a function $\phi \in C^2(\overline{\mathscr{D}})$ such that (1) is satisfied. The following lemma states that no such ϕ exists if $\widehat{V}$ is not differentiable at $X \in \mathscr{D}$ and $(p, A) \in \bar{J}^{2,-}\widehat{V}(X)\backslash J^{2,-}\widehat{V}(X)$.

Lemma 6.13. *Let $\widehat{V} : \overline{\mathscr{D}} \to \mathbb{E}$ be concave and continuous. Suppose $(p, A) \in \bar{J}^{2,-}\widehat{V}(X)\backslash J^{2,-}\widehat{V}(X)$ for some $X \in \mathscr{D}$, and that $\widehat{V}$ is not differentiable at X. Then there are no $\phi_n, \phi \in C^2(\overline{\mathscr{D}}) \cap C_1(\overline{\mathscr{D}})$ and $(p_n, A_n) \in \mathbb{R}^2 \times \mathbb{S}^2$, $n \in \mathbb{N}$, such that conditions (1)-(5) stated above are satisfied.*

Proof. Assume functions ϕ, ϕ_n and a sequence $\{(X_n, p_n, A_n)\}_{n\in\mathbb{N}}$ with the desired properties (1)-(5) exist. We will obtain a contradiction by showing that $\widehat{V}$ must be differentiable at X.

First we will prove that all directional derivatives of $\widehat{V}$ exist, and that the directional derivative in direction $(\cos\theta, \sin\theta)$ is equal to $p_x \cos\theta + p_y \sin\theta$ for all $\theta \in [0, 2\pi)$ (step 1). Then we will prove by contradiction that $\widehat{V}$ is differentiable at (x, y) (step 2). We will obtain a contradiction in step 2 by constructing a sequence of points with some wanted properties converging towards (x, y), we will show that we may assume the points lie approximately on a straight line, and we will obtain a contradiction by using the concavity of $\widehat{V}$. Note that differentiability of $\widehat{V}$ at (x, y) does not follow immediately from the existence of the partial derivatives of $\widehat{V}$ at (x, y). The existence of the partial derivatives of $\widehat{V}$ at (x, y) only implies that $\widehat{V}$ is differentiable at (x, y) if the partial derivatives are *continuous* in a neighbourhood of (x, y), see Proposition 3.33 and Chapter 3.5 of [29].

We can assume without loss of generality that $\widehat{V}(X_n) = \phi_n(X_n)$ for all $n \in \mathbb{N}$. This implies that $\widehat{V}(X) = \phi(X)$. First note that $\phi \leq \widehat{V}$ everywhere; otherwise we would have had $\phi_n > \widehat{V}$ for some $Y \in \overline{\mathscr{D}}$ and all sufficiently large n, which is a contradiction to the assumption that X_n is a global minimum of $\widehat{V} - \phi$ and $(\widehat{V} - \phi)(X_n) = 0$.

Step 1: We will prove that all directional derivatives of $\widehat{V}$ exist, and that the derivative in direction $(\cos\theta, \sin\theta)$ is equal to $p_x \cos\theta + p_y \sin\theta$ for all $\theta \in [0, 2\pi)$. Fix $\theta \in [0, 2\pi)$, and we will show the statement by contradiction. Assume the directional derivative in direction $(\cos\theta, \sin\theta)$ is either non-existent or not equal to $p_x \cos\theta + p_y \sin\theta$. Then there exists a positive sequence $\{h_n\}_{n\in\mathbb{N}}$ converging to 0, such that

$$\lim_{n\to 0} p_n \neq 0,$$

where

$$p_n := \frac{\widehat{V}(x + h_n \cos\theta, y + h_n \sin\theta) - \widehat{V}(x, y) - p_x h_n \cos\theta - p_y h_n \sin\theta}{h_n}. \tag{6.23}$$

By taking a subsequence if necessary, we can assume there is an $\epsilon > 0$ such that we have either (i) $p_n \geq \epsilon$ for all n, or (ii) $p_n \leq -\epsilon$ for all n.

We start to consider case (i). Since $\widehat{V}$ is concave, we have

$$\begin{aligned}\widehat{V}(x{+}h_n \cos\theta, y + h_n \sin\theta) &- \widehat{V}(x, y)\\ &\leq \widehat{V}(x, y) - \widehat{V}(x - h_n \cos\theta, y - h_n \sin\theta)\end{aligned}$$

for all $n \in \mathbb{N}$. Using this inequality, (6.23) and $\widehat{V} \geq \phi$, we get

$$p_n h_n + p_x h_n \cos\theta + p_y h_n \sin\theta \leq \widehat{V}(x, y) - \phi(x - h_n \cos\theta, y - h_n \sin\theta).$$

Taylor expanding ϕ around (x, y) and using $p = D\phi(x, y)$, we get

$$p_n h_n \leq O(h_n^2),$$

and letting $n \to \infty$ we get a contradiction to the assumption $p_n > \epsilon$ for all $n \in \mathbb{N}$.

Now we consider case (ii). We have

$$\begin{aligned}\widehat{V}(x + h_n \cos\theta, y + h_n \sin\theta) &- \phi(x + h_n \cos\theta, y + h_n \sin\theta)\\ &= \big(\widehat{V}(x, y) + p_x h_n \cos\theta + p_y h_n \sin\theta + p_n h_n\big)\\ &\quad - \big(\widehat{V}(x, y) + p_x h_n \cos\theta + p_y h_n \sin\theta + O(h_n^2)\big)\\ &= p_n h_n + O(h_n^2)\end{aligned}$$

by (6.23) and since $p = D\phi(x,y)$. Since $\widehat{V} \geq \phi$ and $p_n < -\epsilon$ for all n, we get a contradiction when we let $n \to \infty$.

We have obtained a contradiction in both case (i) and case (ii). It follows that all directional derivatives of $\widehat{V}$ exist, and that the derivative in direction $(\cos\theta, \sin\theta)$ is equal to $p_x \cos\theta + p_y \sin\theta$.

Step 2: Since $\widehat{V}$ is not differentiable at (x,y), there is a sequence $\{\bar{h}_n\}_{n\in\mathbb{N}}$ with $\bar{h}_n = h_n(\cos\theta_n, \sin\theta_n)$, $h_n \in (0,\infty)$, $\theta \in [0,2\pi]$ and $h_n \to 0$, such that

$$\frac{\widehat{V}\big((x,y)+\bar{h}_n\big) - \widehat{V}(x,y) - p_x h_n \cos\theta_n - p_y h_n \sin\theta_n}{h_n} \nrightarrow 0. \tag{6.24}$$

It is tempting to get a contradiction to (6.24) by writing

$$\widehat{V}\big((x,y)+\bar{h}_n\big) = \widehat{V}(x,y) + p_x h_n \cos\theta_n + p_y h_n \sin\theta_n + O(h_n^2), \tag{6.25}$$

and let $h_n \to 0$, since we know that the partial derivatives of $\widehat{V}$ are p_x and p_y, respectively. However, we only know that (6.25) is valid if there is a $\theta \in [0,2\pi)$ such that $\theta_n = \theta$ for all n, so we need to be more careful when deriving a contradiction.

By taking a subsequence if necessary, we may assume $\theta_n \to \theta$ for some $\theta \in [0,2\pi)$, where θ and $\theta + 2\pi$ are considered to be the same angle for all $\theta \in \mathbb{R}$. We may also assume we have either $\theta_n \geq \theta$ for all n, or $\theta_n \leq \theta$ for all n, and we assume without loss of generality that $\theta_n \leq \theta$ for all $n \in \mathbb{N}$.

Define $\bar{h}'_n := h_n(\cos\theta, \sin\theta)$. We have

$$\frac{\widehat{V}\big((x,y)+\bar{h}'_n\big) - \widehat{V}(x,y) - p_x h_n \cos\theta - p_y h_n \sin\theta}{h_n} \to 0,$$

since the directional derivative in direction $(\cos\theta, \sin\theta)$ is $p_x\cos\theta + p_y \sin\theta$. We obtain a contradiction to (6.24) if we manage to show that

$$\begin{aligned}
&\frac{\widehat{V}\big((x,y)+\bar{h}'_n\big) - \widehat{V}(x,y) - p_x h_n \cos\theta - p_y h_n \sin\theta}{h_n} \\
&\qquad - \frac{\widehat{V}\big((x,y)+\bar{h}_n\big) - \widehat{V}(x,y) - p_x h_n \cos\theta_n - p_y h_n \sin\theta_n}{h_n} \\
&\quad = \frac{\widehat{V}\big((x,y)+\bar{h}'_n\big) - \widehat{V}\big((x,y)+\bar{h}_n\big)}{h_n} + p_x(\cos\theta - \cos\theta_n) + p_y(\sin\theta - \sin\theta_n),
\end{aligned}$$

converges to 0 as $n \to \infty$. We see immediately that the last two terms on the right-hand side converge to 0, so it remains to show that

$$\frac{\widehat{V}\big((x,y)+\bar{h}'_n\big) - \widehat{V}\big((x,y)+\bar{h}_n\big)}{h_n} \to 0. \tag{6.26}$$

We assume the opposite, i.e.,

$$\lim_{n\to\infty} q_n \nrightarrow 0,$$

where

$$q_n := \frac{\widehat{V}\big((x,y)+\bar{h}'_n\big) - \widehat{V}\big((x,y)+\bar{h}_n\big)}{h_n}$$

for all $n \in \mathbb{N}$. By taking a subsequence, we may assume that there is an $\epsilon > 0$ such that (i) $q_n > \epsilon$ for all n, or (ii) $q_n < -\epsilon$ for all n.

First we consider case (i). Fix any $\Delta\theta > 0$. By taking a subsequence, we may assume $\theta_n \in (\theta - \Delta\theta, \theta]$ for all $n \in \mathbb{N}$, and by concavity

$$\frac{\widehat{V}\big((x,y)+\bar{h}_n'\big) - \widehat{V}\big((x,y)+\bar{h}_n\big)}{d_n'} \leq \frac{\widehat{V}\big((x,y)+\bar{h}_n'\big) - \widehat{V}\big((x,y)+\bar{h}_n^{\Delta}\big)}{d_n^{\Delta}},$$

where d_n' is the distance between h_n' and h_n, d_n^{Δ} is the distance between h_n' and h_n^{Δ}, and h_n^{Δ} is the intersection between the line going through h_n' and h_n, and the line going through the origin with slope $\tan(\theta - \Delta\theta)$. We see that $|\bar{h}_n^{\Delta}| \to 0$ as $n \to \infty$.

We know that

$$\widehat{V}\big((x,y)+\bar{h}_n'\big) \leq \widehat{V}(x,y) + p_x h_n \cos\theta + p_y h_n \sin\theta + a' h_n^2$$

and

$$\begin{aligned}\widehat{V}\big((x,y)+\bar{h}_n^{\Delta}\big) \geq\, & \widehat{V}(x,y) + p_x|\bar{h}_n^{\Delta}|\cos(\theta-\Delta\theta) \\ & + p_y|\bar{h}_n^{\Delta}|\sin(\theta-\Delta\theta) - a^{\Delta}|\bar{h^{\Delta}}_n|^2\end{aligned}$$

for some $a', a^{\Delta} > 0$, because we have an expression for the directional derivative in both direction $(\cos\theta, \sin\theta)$ and in direction $\big(\cos(\theta-\Delta\theta), \sin(\theta-\Delta\theta)\big)$. It follows that the right-hand side of (7.35) converges to 0 as $n \to \infty$.

We have $d_n' < h_n(\theta - \theta_n)$ for all $n \in \mathbb{N}$, so

$$\begin{aligned}\frac{\widehat{V}\big((x,y)+h_n'\big) - \widehat{V}\big((x,y)+h_n\big)}{d_n'} &\geq \frac{\widehat{V}\big((x,y)+h_n'\big) - \widehat{V}\big((x,y)+h_n\big)}{h_n'}\frac{1}{\theta-\theta_n} \\ &> \frac{\epsilon}{\theta-\theta_n} \\ &\to \infty.\end{aligned}$$

We have obtained a contradiction.

We obtain a contradiction in the exactly same way in case (ii), except that we define $\bar{h}_n$ to lie on the line through the origin with slope $\tan(\theta + \Delta\theta)$, instead of on the line through the origin with slope $\tan(\theta - \Delta\theta)$.

We have obtained a contradiction in both case (i) and case (ii), so (6.26) holds, and $\widehat{V}$ is differentiable at X. We can conclude that no functions ϕ, ϕ_n and sequences $\{(X_n, p_n, A_n)\}$ with the wanted properties exist. □

The lemma above shows that appropriate functions ϕ and ϕ_n do not exist in general. However, note that functions ϕ and ϕ_n satisfying (1)-(5) always exist for *smooth* $\widehat{V}$. We can simply choose ϕ such that $\widehat{V} - \phi$ has a global minimum at X, and can define $\phi_n = \phi$ and $X_n = X$ for all $n \in \mathbb{N}$.

6.3 Existence and uniqueness

This section contains existence and uniqueness results for the constrained viscosity solution of (5.11) and (4.2). First it is shown that the value function V is a viscosity solution, i.e., existence of a viscosity solution is shown. Then a comparison principle that implies uniqueness is shown.

The theorem below states that V is a viscosity solution of (5.11) and (4.2). Proofs of a similar theorem for the infinite-horizon case is given in [9] and [7]. The proof of the supersolution property of V is similar in the two articles, while the subsolution proof is different. In [9] the subsolution property is shown by introducing a time τ defined as the first time the Lévy process jumps. This approach is not entirely correct, as the time τ is not necessarily well-defined, for example there may be cases where the Lévy process is expected to jump infinitely many times on any time interval. In [7] the subsolution proof is based on the existence of an optimal control, and it is not necessary to introduce a stopping time τ. The proof presented below follows the same idea as the proof in [7].

Theorem 6.14. *The value function V is a viscosity solution of the terminal value problem* (5.11) *and* (4.2).

Proof. First we will prove that V is a supersolution in $\mathscr{D}_T$. Suppose $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$, that $(t,x,y) \in \mathscr{D}_T$ is a global minimizer of $V - \phi$, and that $V(t,x,y) = \phi(t,x,y)$. By Lemma 6.5, V is a supersolution if we can show that

$$\max\left\{G(D_X\phi);\, \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big)\right\} \leq 0.$$

Using Lemma 4.8, we get

$$\phi(t,x,y) = V(t,x,y) \geq V(t,x-c,y+\beta c) \geq \phi(t,x-c,y+\beta c)$$

for all $c \in (0,x]$. Dividing by c and sending $c \to 0$, we get

$$\beta\phi_y(t,x,y) - \phi_x(t,x,y) \leq 0. \tag{6.27}$$

Using the dynamic programming principle and the definition of V with $C \equiv 0$, $\pi \equiv \pi'$, $\pi' \in [0,1]$, we get:

$$V(t,x,y) \geq \mathrm{E}\left[\int_t^{t+\Delta t} e^{-\delta s} U\left(Y_s^{\pi',0}\right) ds + V\left(t+\Delta t, X_{t+\Delta t}^{\pi',0}, Y_{t+\Delta t}^{\pi',0}\right)\right]$$

for all $\Delta t \in [0, T-t]$. Using the fact that $V - \phi$ has a minimum at (t,x,y), we get

$$\phi(t,x,y) \geq \mathrm{E}\left[\int_t^{t+\Delta t} e^{-\delta s} U\left(Y_s^{\pi',0}\right) ds + \phi\left(t+\Delta t, X_{t+\Delta t}^{\pi',0}, Y_{t+\Delta t}^{\pi',0}\right)\right].$$

Applying Itô's formula (Theorem 2.22) on the term $\phi\left(t+\Delta t, X_{t+\Delta t}^{\pi',0}, Y_{t+\Delta t}^{\pi',0}\right)$ on the right-hand side, we get:

$$\begin{aligned}
0 \geq \mathrm{E}\bigg[\int_t^{t+\Delta t} & \phi_t + e^{-\delta s} U\left(Y_s^{\pi',0}\right) + \phi_x\big(\hat{r}(1-\pi_s) + \pi_s\hat{\mu}\big)X_s^{\pi',0} - \beta\phi_y Y_s^{\pi',0} \\
&+ \frac{1}{2}\phi_{xx}\left(\sigma\pi_s X_s^{\pi',0}\right)^2 + \mathscr{I}^{\pi'}\left(s, X_s^{\pi',0}, Y_s^{\pi',0}, \phi\right) ds\bigg] \\
\geq \Delta t \inf_{s\in[t,t+\Delta t]} & \bigg[\phi_t + e^{-\delta s} U\left(Y_s^{\pi',0}\right) + \phi_x\big(\hat{r}(1-\pi_s) + \pi_s\hat{\mu}\big)X_s^{\pi',0} - \beta\phi_y Y_s^{\pi',0} \\
&+ \frac{1}{2}\phi_{xx}\left(\sigma\pi_s X_s^{\pi',0}\right)^2 + \mathscr{I}^{\pi}\left(s, X_s^{\pi',0}, Y_s^{\pi',0}, \phi\right)\bigg].
\end{aligned}$$

We divide this inequality by Δt, and let $\Delta t \to 0$. Since $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ are right-continuous, and ϕ and $\mathscr{I}^\pi$ are smooth, we see that

$$\begin{aligned} 0 \geq \phi_t + e^{-\delta t} U(y) + \phi_x\big(\hat{r}(1-\pi') + \pi'\hat{\mu}\big)x - \beta\phi_y y \\ + \frac{1}{2}\phi_{xx}(\sigma\pi' x)^2 + \mathscr{I}^{\pi'}(t,x,y,\phi). \end{aligned} \tag{6.28}$$

This expression is valid for all $\pi' \in [0,1]$, and therefore

$$\phi_t + F\big(t,x,y,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,x,y,\phi)\big) \leq 0.$$

From (6.27) and (6.28), we see that V is a viscosity supersolution.

We will now prove that V is a subsolution. Suppose ϕ is smooth, and that $(t,x,y) \in \overline{\mathscr{D}_T}$ is a global maximizer of $V-\phi$. By Lemma 6.6, we may assume without loss of generality that $V(t,x,y) = \phi(t,x,y)$, and that there is an $\epsilon' > 0$ such that $(V-\phi)(t',X') \leq (V-\phi)(t,X) - \epsilon'$ for all other maxima $(t',X') \in \overline{\mathscr{D}_T}$.

Arguing by contradiction, we suppose that the subsolution inequality (4.3) is violated. The value function V is continuous by Theorem 4.13, and since ϕ also is continuous, there exists an $\epsilon > 0$ and a non-empty open ball $\mathcal{N}$ centered at (t,x,y) such that

$$\beta\phi_y - \phi_x \leq 0 \tag{6.29}$$

and

$$\begin{aligned} -\epsilon > \phi_t + e^{-\delta t'}U(y') + \beta y'\phi_y + \max_{\pi\in[0,1]}\Big[(\hat{r}(1-\pi) + \pi\hat{\mu})x'\phi_x \\ + \frac{1}{2}(\sigma\pi x')^2\phi_{xx} + \mathscr{I}^\pi(t',X',\phi)\Big] \end{aligned} \tag{6.30}$$

for all $(t',X') \in \mathcal{N} \cap \overline{\mathscr{D}_T}$. We can also assume $(T,x',y') \notin \mathcal{N}$ for any $(x',y') \in \mathscr{D}$, and that

$$V \leq \phi - \epsilon \tag{6.31}$$

on $\partial\mathcal{N} \cap \overline{\mathscr{D}_T}$.

By Lemma 3.6, there exists a consumption strategy $(\pi,C) \in \mathscr{A}_{t,x,y}$ such that

$$V(t,x,y) = \mathrm{E}\left[\int_t^{t+\Delta t} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\,dt + V\left(t+\Delta t, X_{t+\Delta t}^{\pi,C}, Y_{t+\Delta t}^{\pi,C}\right)\right] \tag{6.32}$$

for all $\Delta t \in [0,T-t]$. We define

$$\Delta t = \begin{cases} \begin{array}{c} \frac{1}{2}\inf\Big\{s \in (0,T-t] : \\ \phi\left(t+\Delta t, X_{t+\Delta t}^{\pi,C}, Y_{t+\Delta t}^{\pi,C}\right) \notin \mathcal{N}\Big\} \end{array} & \text{if } \phi\left(t+\Delta t, X_{t+\Delta t}^{\pi,C}, Y_{t+\Delta t}^{\pi,C}\right) \notin \mathcal{N} \\ & \text{for some } \Delta t \in [0,T-t], \\ \frac{1}{2}T & \text{if } \phi\left(t+\Delta t, X_{t+\Delta t}^{\pi,C}, Y_{t+\Delta t}^{\pi,C}\right) \in \mathcal{N} \\ & \text{for all } \Delta t \in [0,T-t]. \end{cases}$$

First we suppose $\Delta t = 0$. The Lévy process jumps almost certainly not at time t, so it must almost certainly be the control that causes the jump out of $\mathcal{N}$. Suppose the jump has size $\Delta C > 0$. Denote the segment between (t,x,y) and $(t,x-\Delta C,t+\beta\Delta C)$

by l, and let (t,x',y') be the intersection between $\mathcal{N}$ and l. We see by (6.29) that ϕ is non-increasing along l in $\mathcal{N}\cap\overline{\mathscr{D}_T}$. Using Lemma 4.8 and (6.31), we get

$$V(t,x,y)=V(t,x',y')\leq\phi(t,x',y')-\epsilon\leq\phi(t,x,y)-\epsilon=V(t,x,y)-\epsilon,$$

which is a contradiction.

Now suppose $\Delta t>0$. Using equation (6.32), that $(V-\phi)(t,x,y)=0$ and that $(V-\phi)\leq 0$ elsewhere, we get

$$\phi(t,x,y)\leq \mathrm{E}\left[\int_t^{t+\Delta t} e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\phi\big(t+\Delta t,X_{t+\Delta t}^{\pi,C},Y_{t+\Delta t}^{\pi,C}\big)\right].$$

Using Itô's formula and then (6.29) and (6.30), we get

$$\begin{aligned}0\leq &\mathrm{E}\bigg[\int_t^{t+\Delta t}\phi_t+e^{-\delta s}U(Y_s^{\pi,C})-\phi_x\big(\hat{r}+\pi(\hat{\mu}-\hat{r})\big)X_s^{\pi,C}-\phi_y\beta Y_s^{\pi,C}\\&+\frac{1}{2}(\sigma\pi X_s^{\pi,C})^2\phi_{xx}+\mathscr{I}^\pi(t,x,y,\phi)\,ds+\int_t^{t+\Delta t}-\phi_x+\beta\phi_y dC_s^c\\&+\sum_{\Delta C_s\neq 0}\big(\phi(s,X_s^{\pi,C},Y_s^{\pi,C})-\phi(s,X_s^{\pi,C}+\Delta C,Y_s^{\pi,C}-\beta\Delta C)\big)\bigg]\\\leq&-\epsilon\Delta t,\end{aligned}$$

where C^c denotes the continuous part of C. This inequality contradicts itself, and therefore we see that V also is a subsolution of (5.11) on $\overline{\mathscr{D}_T}$.

It is explained in Section 4.1 that V satisfies (4.2). By Theorem 4.13, $V\in C(\mathscr{D}_T)$, and by Definition 6.2 we see that V is a viscosity solution of (5.11) and (4.2). □

The theorem below shows that subsolutions of (5.11) and (4.2) are less than or equal to supersolutions. The main idea of the proof is inspired by [9]. Here the authors show the same theorem for a infinite-horizon problem in the case of a pure jump Lévy process. It is necessary to do some modifications in the proof because we are considering general Lévy processes, see the discussion in Section 6.2. It is also necessary to do some modifications in the proof because of the finite time-horizon, and these modifications have been inspired by Theorem 9.1, Chapter II, in [25]. Here the authors show a similar theorem in the case of a first-order deterministic control problem on a bounded domain.

Theorem 6.15 (Comparison principle). *Assume $\underline{v}\in C_{\gamma^*}(\overline{\mathscr{D}_T})$ is a subsolution of (5.11) in $[0,T)\times\overline{\mathscr{D}}$, that $\overline{v}\in C_{\gamma^*}(\overline{\mathscr{D}_T})$ is a supersolution of (5.11) in $\mathscr{D}_T$, and that $\underline{v}\leq\overline{v}$ for $t=T$. Then $\underline{v}\leq\overline{v}$ everywhere in $\overline{\mathscr{D}_T}$.*

Proof. Since $\underline{v},\overline{v}\in C_{\gamma^*}(\overline{\mathscr{D}_T})$, there is a $\gamma'>0$ such that $\delta>k(\gamma')$ and $\underline{v},\overline{v}\in C_{\gamma'}(\overline{\mathscr{D}_T})$.

Let w be defined as in Lemma 6.10, and choose $\tilde{K}$ so large that

$$\overline{v}^\theta=(1-\theta)\overline{v}+\theta w$$

is a strict supersolution of (5.11) on any bounded domain of $\mathscr{D}_T$, and such that $w>\underline{v}$ for $t=T$. Such a value of $\tilde{K}$ exists by Lemma 6.10, and because $\bar{\gamma}>\gamma'$. Instead of comparing $\overline{v}$ and $\underline{v}$, we will compare $\overline{v}^\theta$ and $\underline{v}$. By letting $\theta\to 0$, we will obtain a comparison result for $\overline{v}$ and $\underline{v}$. We know that

$$\underline{v}(t,X) - \overline{v}^{\theta}(t,X) \to -\infty$$

as $x, y \to \infty$, and that

$$\underline{v}(t,X) - \overline{v}^{\theta}(t,X) < 0$$

for $t = T$. When finding the maximum of $\underline{v}(t,X) - \overline{v}^{\theta}(t,X)$, it is therefore sufficient to consider the domain

$$\mathscr{O}_T := [0,T) \times \mathscr{O},$$

where

$$\mathscr{O} = \{(x,y) \in \mathscr{D} : 0 < x, y < R\}.$$

for some $R > 0$. Assume that $\underline{v} \le \overline{v}^{\theta}$ is violated somewhere in $\mathscr{O}_T$, and define

$$M := \max_{\overline{\mathscr{O}_T}} (\underline{v} - \overline{v}^{\theta})(t,X) + \xi(t-T),$$

where $\xi > 0$ is chosen so small that $M > 0$. Assume $(t^*, X^*) \in \mathscr{O}_T$ satisfies

$$M = (\underline{v} - \overline{v}^{\theta})(t^*, X^*) + \xi(t^* - T). \tag{6.33}$$

By choosing R large enough, we may assume that we have either

1. $(t^*, X^*) \in \Gamma$, or
2. $(t^*, X^*) \in \mathscr{O}_T$,

where $\Gamma = \{(t,X) \in \overline{\mathscr{O}_T} : x = 0, y \in [0,R) \text{ or } x \in [0,R), y = 0\}$. We cannot have $t^* = T$, since $\underline{v} \le \overline{v}^{\theta}$ for $t = T$.

Case 1: This is the case $(t^*, X^*) \in \Gamma$. Since $\partial\mathscr{O}_T$ is piecewise linear, there exist constants $h_0, \kappa_0 > 0$ and a uniformly continuous map $\eta : \overline{\mathscr{O}_T} \to \mathbb{R}^3$ satisfying

$$\mathscr{N}\big((t,X) + h\eta(t,X), h\kappa_0\big) \subset \mathscr{O}_T \tag{6.34}$$

for all $(t,X) \in \overline{\mathscr{O}_T}$ and $h \in (0, h_0]$, where $\mathscr{N}(z,\rho)$ denotes the ball in $\mathbb{R}^3$ with radius ρ and centre $z \in \mathbb{R}^3$.

For any $\alpha > 1$ and $\epsilon \in (0,1)$, define the functions ϕ and Φ on $\overline{\mathscr{O}_T} \times \overline{\mathscr{O}_T}$ by

$$\begin{aligned}\phi\big((t_1,X_1),(t_2,X_2)\big) := \Big|\alpha\big((t_1,X_1) - (t_2,X_2)\big) + \epsilon\eta(t^*,X^*)\Big|^2 \\ + \epsilon|(t_1,X_1) - (t^*,X^*)|^2 - \xi(t_2 - T)\end{aligned}$$

and

$$\Phi\big((t_1,X_1),(t_2,X_2)\big) := \underline{v}(t_1,X_1) - \overline{v}^{\theta}(t_2,X_2) - \phi\big((t_1,X_1),(t_2,X_2)\big)$$

Let

$$M_\alpha := \max_{\overline{\mathscr{O}_T} \times \overline{\mathscr{O}_T}} \Phi\big((t_1,X_1),(t_2,X_2)\big).$$

We have $M_\alpha > 0$ for any $\alpha > 1$ and $\epsilon \le \epsilon_0$, where $\epsilon_0 > 0$ is some fixed small number. Let $\big((t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha})\big) \in \overline{\mathscr{O}_T} \times \overline{\mathscr{O}_T}$ be a maximizer of Φ, i.e., $M_\alpha = \Phi\big((t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha})\big)$. By (6.34), we assume that α is so large that $(t^*, X^*) + \frac{\epsilon}{\alpha}\eta(t^*, X^*) \in \mathscr{O}_T$. Using

$$\Phi\big((t_{\alpha,1}, X_{\alpha,1}), (t_{\alpha,2}, Y_{\alpha,2})\big) \ge \Phi\left((t^*, X^*), (t^*, X^*) + \frac{\epsilon}{\alpha}\eta(t^*, X^*)\right),$$

we get

$$\begin{aligned}
&\Big|\alpha\big((t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})\big)+\epsilon\eta(t^*,X^*)\Big|^2+\epsilon|(t_{1,\epsilon},X_{1,\alpha})-(t^*,X^*)|^2\\
&\quad\le \underline{v}(t_{1,\alpha},X_{1,\alpha})-\overline{v}^\theta(t_{2,\alpha},X_{2,\alpha})-\underline{v}(t^*,X^*) \qquad (6.35)\\
&\qquad+\overline{v}^\theta\left((t^*,X^*)+\frac{\epsilon}{\alpha}\eta(t^*,X^*)\right)+\xi\left(t_{2,\alpha}-t^*-\frac{\epsilon}{\alpha}\eta(t^*,X^*)\right).
\end{aligned}$$

The right-hand side of this equation is bounded as $\alpha\to\infty$, and therefore $\alpha|(t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})|$ is bounded uniformly in α. It follows that

$$\lim_{\alpha\to\infty}\Big((t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})\Big)=0$$

and

$$\lim_{\alpha\to\infty}\Big(\underline{v}(t_{1,\alpha},X_{1,\alpha})-\overline{v}^\theta(t_{2,\alpha},X_{2,\alpha})\Big)\le M.$$

Sending $\alpha\to\infty$ in (6.35) and using the uniform continuity of $\underline{v}$ and $\overline{v}^\theta$ in $\overline{\mathscr{O}_T}$ we see that $\alpha\big((t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})\big)+\epsilon\eta(t^*,X^*)\to 0$, $(t_{1,\alpha},X_{1,\alpha}),(t_{2,\alpha},X_{2,\alpha})\to(t^*,X^*)$ and $M_\alpha\to M$ as $\alpha\to\infty$. Therefore, using the uniform continuity of η,

$$\begin{aligned}
(t_{2,\alpha},X_{2,\alpha})&=(t_{1,\alpha},X_{1,\alpha})+\frac{\epsilon}{\alpha}\eta(t^*,X^*)+O\left(\frac{1}{\alpha}\right)\\
&=(t_{1,\alpha},X_{1,\alpha})+\frac{\epsilon}{\alpha}\eta(t_{1,\alpha},X_{1,\alpha})+O\left(\frac{1}{\alpha}\right),
\end{aligned}$$

and we use (6.34) to get $(t_{2,\alpha},X_{2,\alpha})\in\mathscr{O}_T$ for α large enough.

We have

$$D^2\phi=2\alpha\begin{pmatrix}I&-I\\-I&I\end{pmatrix}+2\epsilon\begin{pmatrix}I&0\\0&0\end{pmatrix}.$$

We see that the conditions of Lemma 6.12 are satisfied, so, for any $\varsigma\in(0,1)$, there exist matrices $A_1,A_2\in\mathbb{S}^3$ such that

$$\begin{pmatrix}A_1&0\\0&-A_2\end{pmatrix}\le 2\frac{\alpha}{\varsigma}\begin{pmatrix}I&I\\-I&I\end{pmatrix}+2\epsilon\begin{pmatrix}I&0\\0&0\end{pmatrix}, \qquad (6.36)$$

$$\begin{aligned}
0\le\max\big\{&G(D_{X_1}\phi);\ \phi_{t_1}+F\big(t_{1,\alpha},X^*_{1,\alpha},D_{X_1}\phi,\widehat{A}_1,\\
&\mathscr{I}^{\pi,\kappa}(t_{1,\alpha},X_{1,\alpha},\underline{v},D_{X_1}\phi),\mathscr{I}^{\pi}(t_{1,\alpha},X_{1,\alpha},\phi)\big)\big\}, \qquad (6.37)
\end{aligned}$$

and

$$\begin{aligned}
-\theta f\ge\max\big\{&G(-D_{X_2}\phi);\ -\phi_{t_2}+F\big(t_{2,\alpha},X_{2,\alpha},-D_{X_2}\phi,\widehat{A}_2,\\
&\mathscr{I}^{\pi,\kappa}(t_{2,\alpha},X_{2,\alpha},\overline{v}^\theta,-D_{X_2}\phi),\mathscr{I}^{\pi}_\kappa(t_{2,\alpha},X_{2,\alpha},-\phi)\big)\big\}, \qquad (6.38)
\end{aligned}$$

where ϕ_t and $D_X\phi$ are evaluated at $\big((t_{1,\alpha},X_{1,\alpha}),(t_{2,\alpha},X_{2,\alpha})\big)$, and $\widehat{A}_1$ and $\widehat{A}_2$ are the parts of A_1 and A_2 corresponding to X_1 and X_2, respectively.

We know that $\widehat{A}_i$ is a two-dimensional matrix, and we denote its elements by $a_{i,xx}$, $a_{i,xy}$, etc. Multiplying (6.36) by $(x_{1,\alpha}e_2^t\ x_{2,\alpha}e_2^t)$ on the left hand side and $(x_{1,\alpha}e_2^t\ x_{2,\alpha}e_2^t)^t$ on the right-hand side, where $e_2=(0,1,0)^t\in\mathbb{R}^3$ and t denotes the transpose of a matrix, we get

$$x_{1,\alpha}^2a_{1,xx}-x_{2,\alpha}^2a_{2,xx}\le\frac{\alpha}{\varsigma}(x_{1,\alpha}-x_{2,\alpha})^2+\epsilon x_{1,\alpha}^2,$$

so

$$\lim_{\epsilon\to0}\lim_{\alpha\to\infty}\big(x_{1,\alpha}^2a_{1,xx}-x_{2,\alpha}^2a_{2,xx}\big)\le 0.$$

Equation (6.38) implies that $G(-D_{X_2}\phi) \le -\theta f$ and

$$-\theta f \ge -\phi_{t_2} + F\big(t_{2,\alpha}, X_{2,\alpha}, -D_{X_2}\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \overline{v}^\theta, -D_{X_2}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big). \tag{6.39}$$

Next we claim that $G(D_{X_1}\phi) < 0$ for sufficiently large α. Assume to the contrary that $G(D_{X_1}\phi) \ge 0$. Then it follows that

$$\begin{aligned}
-\theta f &\ge G(-D_{X_2}\phi) - G(D_{X_1}\phi)\\
&= -\beta(\phi_{y_2} + \phi_{y_1}) + (\phi_{x_2} + \phi_{x_1})\\
&= -2\beta\epsilon(y_{1,\alpha} - y^*) + 2\epsilon(x_{1,\alpha} - x^*),
\end{aligned}$$

which converges to zero as $\alpha \to \infty$. This is a contradiction to the fact that f is strictly positive on $\mathscr{O}_T$. Thus our claim holds. Equation (6.37) and $G(D_{X_1}\phi) < 0$ implies that

$$\phi_{t_1} + F\big(t_{1,\alpha}, X_{1,\alpha}, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi), \mathscr{I}^{\pi}_{\kappa}(X_{1,\alpha}, \phi)\big) \ge 0. \tag{6.40}$$

Using (6.39) and (6.40), we get

$$\begin{aligned}
\theta f \le\; &\phi_{t_1} + F\big(t_{1,\alpha}, X_{1,\alpha}, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi)\big)\\
&+ \phi_{t_2} - F\big(t_{2,\alpha}, X_{2,\alpha}, -D_{X_2}\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \overline{v}^\theta, -D_{X_2}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big)\\
\le\; &\big(e^{-\delta t_{1,\alpha}}U(y_{1,\alpha}) - e^{-\delta t_{2,\alpha}}U(y_{2,\alpha})\big) + \big(\phi_{t_1} + \phi_{t_2}\big) - \beta\big(y_{1,\alpha}\phi_{y_1} + y_{2,\alpha}\phi_{y_2}\big)\\
&+ \max_{\pi\in[0,1]}\bigg[\big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big)(x_{1,\alpha}\phi_{x_1} + x_{2,\alpha}\phi_{x_2})\\
&+ \frac{1}{2}(\sigma\pi)^2(x_{1,\alpha}^2 a_{1,\alpha,xx} - x_{2,\alpha}^2 a_{2,\alpha,xx})\\
&+ \big(\mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v}, -D_{X_2}\phi)\big)\\
&+ \big(\mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big)\bigg]\\
\le\; &\big(e^{-\delta t_{1,\alpha}}U(y_{1,\alpha}) - e^{-\delta t_{2,\alpha}}U(y_{2,\alpha})\big) + 2\epsilon t_{1,\alpha} - \xi - \beta\epsilon y_{1,\alpha}(y_{1,\alpha} - y^*)\\
&+ \max_{\pi\in[0,1]}\bigg[2\big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big)\epsilon x_{1,\alpha}(x_{1,\alpha} - x^*)\\
&+ \frac{1}{2}(\sigma\pi)^2(x_{1,\alpha}^2 a_{1,\alpha,xx} - x_{2,\alpha}^2 a_{2,\alpha,xx})\\
&+ \big(\mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v}, -D_{X_2}\phi)\big)\\
&+ \big(\mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big)\bigg].
\end{aligned} \tag{6.41}$$

It is shown in [9] that

$$\lim_{\epsilon\to 0}\lim_{\alpha\to\infty} \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v}, D_{X_2}\phi) \le 0,$$

and the convergence is uniform in $\kappa > 0$. We also know that

$$\mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, \phi) \to 0$$

as $\kappa \to 0$. By letting, in that order, $\alpha \to \infty$, $\epsilon \to 0$ and $\kappa \to 0$, we see that right-hand side of (6.41) converges to something < 0. This is a contradiction to (6.41).

Case 2: We will now consider Case 2. For any $\alpha > 1$ and $\epsilon \in (0,1)$ define functions $\phi, \Phi(X,Y) : \overline{\mathscr{O}_T} \times \overline{\mathscr{O}_T} \to \mathbb{R}$ by

$$\phi\big((t_1,X_1),(t_2,X_2)\big) := \frac{\alpha}{2}|(t_1,X_1)-(t_2,X_2)|^2 - \xi(t_2-T),$$

$$\Phi\big((t_1,X_1),(t_2,X_2)\big) := \underline{v}(t_1,X_1) - \overline{v}^\theta(t_2,X_2) - \phi\big((t_1,X_1),(t_2,X_2)\big).$$

Let

$$M_\alpha = \max_{\overline{\mathscr{O}_T}\times\overline{\mathscr{O}_T}} \Phi\big((t_1,X_1),(t_2,X_2)\big).$$

We have $M_\alpha \geq M > 0$ for all $\alpha > 1$. Let $\big((t_{1,\alpha},X_{1,\alpha}),(t_{2,\alpha},X_{2,\alpha})\big)$ be a maximizer of Φ, i.e., $M_\alpha = \Phi\big((t_{1,\alpha},X_\alpha),(t_{2,\alpha},X_{2,\alpha})\big)$, and note that the inequality

$$\Phi\big((t_{1,\alpha},X_{1,\alpha}),(t_{1,\alpha},X_{1,\alpha})\big) + \Phi\big((t_{2,\alpha},X_{2,\alpha}),(t_{2,\alpha},X_{2,\alpha})\big) \leq 2\Phi\big((t_{1,\alpha},X_{1,\alpha}),(t_{2,\alpha},X_{2,\alpha})\big)$$

implies

$$\begin{aligned}\alpha|&(t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})|^2\\ &\leq \underline{v}(t_{1,\alpha},X_{1,\alpha}) - \underline{v}(t_{2,\alpha},X_{2,\alpha}) + \overline{v}^\theta(t_{1,\alpha},X_{1,\alpha}) - \overline{v}^\theta(t_{2,\alpha},X_{2,\alpha})\\ &\quad + \xi(t_{2,\alpha}-t_{1,\alpha}).\end{aligned}$$

The right-hand side of this equation is bounded as $\alpha \to \infty$, as all the functions are uniformly continuous on $\overline{\mathscr{O}_T}\times\overline{\mathscr{O}_T}$. Therefore we see that $|(t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})| \to 0$ as $\alpha \to \infty$. Using the uniform continuity of $\underline{v}, \overline{v}^\theta$ in $\mathscr{O}_T$ again, we see that $\alpha|(t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})| \to 0$ as $\alpha \to \infty$. Now we use $M \leq M_\alpha$ and (6.33) to get

$$\begin{aligned}0 &= \lim_{\alpha\to\infty} \frac{\alpha}{2}\big|(t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})\big|^2\\ &\leq \lim_{\alpha\to\infty} \underline{v}(t_{1,\alpha},X_{1,\alpha}) - \overline{v}^\theta(t_{2,\alpha},X_{2,\alpha}) + \xi(t_{2,\alpha}-T) - M\\ &\leq 0,\end{aligned}$$

and we conclude that $M_\alpha \to M$ as $\alpha \to \infty$. Since $M > 0$ and $\underline{v} \leq \overline{v}^\theta$ on Γ, we see that any limit point of $\big((t_{1,\alpha},X_{1,\alpha}),(t_{2,\alpha},X_{2,\alpha})\big)$ belongs to $\mathscr{O}_T \times \mathscr{O}_T$. For large enough α, we therefore have $(t_{1,\alpha},X_{1,\alpha}),(t_{2,\alpha},X_{2,\alpha}) \in \mathscr{O}_T$.

We have

$$D^2\phi = \alpha\begin{pmatrix} I & -I\\ -I & I\end{pmatrix},$$

so by defining $g_0 \equiv \alpha$, $g_1 \equiv 0$ and $g_2 \equiv 0$, we see that Lemma 6.12 can be applied. It follows that, for any $\varsigma \in (0,1)$, there exist matrices $A_1, A_2 \in \mathbb{S}^3$ such that

$$\begin{pmatrix} A_1 & 0\\ 0 & -A_2\end{pmatrix} \leq \frac{\alpha}{\varsigma}\begin{pmatrix} I & -I\\ -I & I\end{pmatrix},$$

$$\begin{aligned}0 \leq \max\big\{&G(D_{X_1}\phi);\ \phi_{t_1} + F\big(t_1^*, X_1^*, D_{X_1}\phi, \widehat{A}_1,\\ &\mathscr{I}^{\pi,\kappa}(t_1^*, X_1^*, \underline{v}, D_{X_1}\phi), \mathscr{I}^\pi(t_1^*, X_1^*, \phi)\big)\big\},\end{aligned} \tag{6.42}$$

and

$$\begin{aligned}-\theta f \geq \max\big\{&G(-D_{X_2}\phi);\ -\phi_{t_2} + F\big(t_2^*, X_2^*, -D_{X_2}\phi, \widehat{A}_2,\\ &\mathscr{I}^{\pi,\kappa}(t_2^*, X_2^*, \overline{v}^\theta, -D_{X_2}\phi), \mathscr{I}_\kappa^\pi(t_2^*, X_2^*, -\phi)\big)\big\}.\end{aligned}$$

As in Case 1, we get

$$\lim_{\alpha\to\infty}\left(x_{1,\alpha}^2 a_{1,xx} - x_{2,\alpha}^2 a_{2,xx}\right) \le 0,$$

$$G(-D_{X_2}\phi) \le -\theta f$$

and

$$-\phi_{t_2} + F\big(t_2^*, X_2^*, -D_{X_2}\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_2^*, X_2^*, \overline{v}^\theta, -D_{X_2}\phi), \mathscr{I}_\kappa^\pi(t_2^*, X_2^*, -\phi)\big) \le -\theta f. \tag{6.43}$$

We have $G(D_{X_1}\phi) < 0$, because $G(D_{X_1}\phi) \equiv G(-D_{X_2}\phi) < 0$. Using (6.42), we get

$$\phi_{t_1} + F\big(t_1^*, X_1^*, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_1^*, X_1^*, \underline{v}, D_{X_1}\phi), \mathscr{I}_\kappa^\pi(t_1^*, X_1^*, \phi)\big) \ge 0. \tag{6.44}$$

Using (6.43) and (6.44), we get

$$\begin{aligned}
\theta f \le{}& \phi_{t_1} + F\big(t_1^*, X_1^*, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_1^*, X_1^*, \underline{v}, D_{X_1}\phi), \mathscr{I}_\kappa^\pi(t_1^*, X_1^*, \phi)\big)\\
&\phi_{t_2} - F\big(t_2^*, X_2^*, -D_{X_2}\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_2^*, X_2^*, \overline{v}^\theta, -D_{X_2}\phi), \mathscr{I}_\kappa^\pi(t_2^*, X_2^*, -\phi)\big)\\
&\le \big(e^{-\delta t_{1,\alpha}} U(y_{1,\alpha}) - e^{-\delta t_{2,\alpha}} U(y_{2,\alpha})\big) - \xi\\
&\quad + \max_{\pi\in[0,1]}\Bigg[\frac{1}{2}(\sigma\pi)^2(x_{1,\alpha}^2 a_{1,\alpha,xx} - x_{2,\alpha}^2 a_{2,\alpha,xx})\\
&\quad + \Big(\mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v}, -D_{X_2}\phi)\Big)\\
&\quad + \Big(\mathscr{I}_\kappa^\pi(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}_\kappa^\pi(t_{2,\alpha}, X_{2,\alpha}, -\phi)\Big)\Bigg].
\end{aligned} \tag{6.45}$$

As in Case 1, we have

$$\lim_{\epsilon\to 0}\lim_{\alpha\to\infty} \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v}^\theta, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\underline{v}, D_{X_2}\phi) \le 0$$

and

$$\lim_{\kappa\to 0} \mathscr{I}_\kappa^\pi(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}_\kappa^\pi(t_{2,\alpha}, X_{2,\alpha}, -\phi) = 0,$$

so by letting, in that order, $\alpha \to \infty$, $\kappa \to \infty$ in (6.45) we obtain a contradiction. This concludes the proof of the theorem. □

Uniqueness of viscosity solutions follows directly from the comparison principle.

Theorem 6.16 (Uniqueness of viscosity solutions in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$). *Viscosity solutions of the terminal value problem* (5.11) *and* (4.2) *in* $C'_{\gamma^*}(\overline{\mathscr{D}_T})$ *are unique.*

Proof. Let $v_1, v_2 \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$ be two viscosity solutions of (5.11) and (4.2) on $\overline{\mathscr{D}_T}$, i.e., $v_1, v_2 \in C_{\gamma'}(\overline{\mathscr{D}_T})$ for some $\gamma' < \gamma^*$ with $\delta > k(\gamma')$. We know that v_1 is a viscosity subsolution of (5.11) on $\overline{\mathscr{D}_T}$, that v_2 is a supersolution of (5.11) on $\mathscr{D}_T$, and that $v_1 = v_2$ for $t = T$. By Theorem 6.15, we get $v_1 \le v_2$. We have $v_2 \le v_1$ by a similar argument, so $v_1 \equiv v_2$, and uniqueness follows. □

Chapter 7
Explicit solution formulas

In this chapter we will derive explicit solution formulas for two special cases:

1. The Lévy process has only negative jumps, U is a CRRA utility function, and W is given by a specific formula.
2. A generalization of the Merton problem, where the risky asset follows a geometric Lévy process, and the utility functions are of CRRA type.

See Section 2.3 and Section 7.2 for a short description of the Merton problem, and see Section 2.3 for the definition of a CRRA utility function. It will be shown that the functions we find satisfy the HJB equation in a classical sense, and by using Lemma 6.7 we conclude that we have found the correct value functions.

Both formulas are inspired by formulas found in [8] for the infinite-horizon case of our problem. However, the function W, which is not present in the infinite-horizon problem, creates a larger family of solutions in our case. In the first case we consider, an explicit solution formula is found for only one specific function W. In the second case, we succeed in finding explicit solution formulas for a larger class of functions W.

7.1 CRRA utility and only negative jumps

In [8] the authors construct an explicit solution formula and allocation strategy in an infinite-horizon case. It is proved that the solution formula is valid provided some additional constraints on the parameters are satisfied. We will first describe the solution found in [8], and then we will prove that there are no parameter values satisfying all the constraints given in [8]. However, as we will prove below, one of the constraints given in [8] can be replaced by a weaker constraint, and the new set of constraints has solutions. We will construct a solution formula for a finite-horizon problem based on the solution found in [8]. This explicit solution formula will be used in Chapter 15, when we want to analyse the error of our numerically simulated solution.

In infinite-horizon problem described in [8], the utility function U is on the form

$$U(y) = \frac{y^\gamma}{\gamma} \tag{7.1}$$

for some $\gamma \in (0, 1)$, and only negative jumps in the Lévy process are allowed, i.e., $\nu\big((0,\infty)\big) = 0$. The value function $\widehat{V} : \overline{\mathscr{D}} \to [0,\infty)$ is defined by

$$\widehat{V}(x,y) = \sup_{(\pi,C)\in\widehat{\mathscr{A}}_{x,y}} \mathrm{E}\left[\int_0^\infty e^{-\delta t} U(Y_s^{\pi,C})\,ds\right], \tag{7.2}$$

where

$$X_s^{\pi,C} = x - C_s + \int_t^s \big(\hat{r} + (\hat{\mu}-\hat{r})\pi_{s'}\big) X_{s'}^{\pi,C}\,ds' + \int_t^s \sigma \pi_{s'} X_{s'}^{\pi,C}\,dB_{s'}$$
$$+ \int_t^s \pi_{s'-} X_{s'-}^{\pi,C} \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\,\tilde{N}(ds,dz)$$

and

$$Y_s^{\pi,C} = ye^{-\beta(s-t)} + \beta e^{-\beta s}\int_t^s e^{\beta s'}\,dC_{s'}.$$

The initial values of $X^{\pi,C}$ and $Y^{\pi,C}$ are x and y, respectively, and B_t, ν, $\widehat{\mathscr{A}_{x,y}}$ and all the constants involved, satisfy the same conditions as in our problem.

The corresponding HJB equation is

$$\max\left\{\widehat{G}(D_X\hat{v});\ \widehat{F}\big(X,\hat{v},D_X\hat{v},D_X^2\hat{v},\widehat{\mathscr{I}^\pi}(X,\hat{v})\big)\right\}, \tag{7.3}$$

where

$$\widehat{G}(D_X\hat{v}) = \beta\hat{v}_y - \hat{v}_x,$$

$$\begin{aligned}\widehat{F}\big(X,&\hat{v},D_X\hat{v},D_X^2\hat{v},\widehat{\mathscr{I}^\pi}(X,\hat{v})\big)\\ &= U(y) - \delta\hat{v} - \beta y\hat{v}_y + \max_{\pi\in[0,1]}\left[(\hat{r}+(\hat{\mu}-\hat{r})\pi)x\hat{v}_x + \frac{1}{2}(\sigma\pi x)^2\hat{v}_{xx} + \widehat{\mathscr{I}^\pi}(X,\hat{v})\right]\end{aligned}$$

and

$$\widehat{\mathscr{I}^\pi}(X,\hat{v}) = \int_{\mathbb{R}\setminus\{0\}} \hat{v}\big(x+\pi x(e^z-1),y\big) - \hat{v}(x,y) - \pi x\hat{v}_x(x,y)(e^z-1)\,\nu(dz).$$

The authors of [8] show that the value function can be written as

$$\widehat{V}(x,y) = \begin{cases} k_1 y^\gamma + k_2 y^\gamma\left[\frac{x}{ky}\right]^\rho & \text{if } 0 \le x < ky,\\ k_3\left(\frac{y+\beta x}{1+\beta k}\right)^\gamma & \text{if } x \ge ky \ge 0,\end{cases} \tag{7.4}$$

where

$$k_1 = \frac{1}{\gamma(\delta+\beta\gamma)},\ k_2 = \frac{1-\rho}{(\rho-\gamma)(\delta+\beta\gamma)},\ k_3 = \frac{\rho(1-\gamma)}{\gamma(\rho-\gamma)(\delta+\beta\gamma)},\ k = \frac{1-\rho}{\beta(\rho-\gamma)} \tag{7.5}$$

and

$$\begin{aligned}&(\hat{\mu}-\hat{r}) - (1-\rho)\sigma^2\pi^* + \int_{-\infty}^{0^-}\big(1+\pi^*(e^z-1)\big)^{\rho-1}(e^z-1) - (e^z-1)\,\nu(dz) = 0,\\ &\left(\hat{r} + (\hat{\mu}-\hat{r})\pi^* + \beta - \frac{1}{2}(\sigma\pi^*)^2(1-\rho)\right)\rho\\ &\qquad = \delta + \beta\gamma - \int_{-\infty}^{0^-}\big((1+\pi^*(e^z-1))^\rho - 1 - \rho\pi^*(e^z-1)\big)\,\nu(dz),\end{aligned} \tag{7.6}$$

provided (7.6) has a solution (ρ,π^*) satisfying

$$\rho\in(\gamma,1],\ \pi^*\in(0,1). \tag{7.7}$$

The only additional condition on the constants is that this equation is satisfied:

$$\frac{\rho(1-\gamma)}{\rho-\gamma} \geq \frac{\delta+\beta\gamma}{\delta-k_0(\gamma)}. \tag{7.8}$$

It is proved that $\widehat{V}$ satisfies

$$\widehat{G}(X, D_X\hat{v}) = 0$$

for $x \geq ky$ and

$$\widehat{F}\big(X, \hat{v}, D_X\hat{v}, D_X^2\hat{v}, \widehat{\mathscr{I}^\pi}(X,\hat{v})\big) = 0 \tag{7.9}$$

for $x \leq ky$. It is also shown that it is optimal to let a constant fraction π^* of the wealth be invested in the stock, and that an optimal consumption process C_t^* is given by

$$C_t^* = \Delta C_0^* + \int_0^t \frac{X_t^*}{1+\beta k}\, dZ_s,$$

where

$$\Delta C_0^* = \left[\frac{x-kY_{0-}}{1+\beta k}\right]^+,$$

$$Z_t = \sup_{0\leq s\leq t}\left[\ln\frac{\widehat{X}_s}{\widehat{Y}_s} - \ln k\right]^+,$$

$$\widehat{Y}_t = (Y_0 + \beta\Delta C_0^*)e^{-\beta t},$$

and

$$\begin{aligned}\widehat{X}_t = (x-\Delta C_0^*) + \int_0^t \big(\hat{r} + (\hat{\mu}-\hat{r})\pi^*\big)\widehat{X}_s\, ds + \int_0^t \sigma\pi^*\widehat{X}_s\, dB_s \\ + \int_0^t \pi^*\widehat{X}_{s-}\int_{-\infty}^{0-}(e^z-1)\,\tilde{N}(ds,dz).\end{aligned}$$

Our first result is that there are no parameter values satisfying all the constraints mentioned above.

Lemma 7.1. *There are no parameter values satisfying* (7.6)-(7.8) *and the constraints stated in Chapter 3.*

Proof. It is shown in [8] that (7.8) implies

$$\widehat{F}\big(X, \widehat{V}, \widehat{V}_X, \widehat{V}_X^2, \mathscr{I}^\pi(X,\widehat{V})\big) \leq 0,$$

for all $x \geq ky$, where $\widehat{V}$ is given by the explicit formula (7.4). The inequality is strict, unless $\nu \equiv 0$ and $y \equiv 0$, i.e., we have

$$\widehat{F}\big(X, \widehat{V}, \widehat{V}_X, \widehat{V}_X^2, \mathscr{I}^\pi(X,\widehat{V})\big) < 0 \tag{7.10}$$

for $x \geq ky$, unless $\nu \equiv 0$ and $y \equiv 0$. See the derivation on page 460 in [8] for a proof of this. However, since $\widehat{F}\big(X, \widehat{V}, \widehat{V}_X, \widehat{V}_X^2, \mathscr{I}^\pi(X,\widehat{V})\big) = 0$ for $x \leq ky$ and $\widehat{V}$ is smooth at the boundary $x = ky$, we have

$$\widehat{F}\big(X, \widehat{V}, \widehat{V}_X, \widehat{V}_X^2, \mathscr{I}^\pi(X,\widehat{V})\big) = 0$$

for $x = ky$. This is a contradiction to (7.10), and we see that there are no parameter values satisfying all the constraints of the problem. □

We will give an alternative proof of Lemma 7.1 for two special cases. First we will prove Lemma 7.1 for the case $\nu \equiv 0$, and then we will prove the lemma for $\nu(dz) = \lambda\delta_{-\xi}(dz)$, $\sigma = 0$ and $\xi \in (0,1)$, where $\delta_{-\xi}$ is the Dirac measure located at $-\xi$. If $\nu \equiv 0$, the risky asset follows a geometric Brownian, and if $\nu = \lambda\delta_{-\xi}(dz)$, $\sigma = 0$ and $\xi \in (0,1)$, the Lévy process can be written as $L_s = \mu s + \xi N_t$, where N_t is a Poisson process.

Alternative proof of Lemma 7.1 for $\nu \equiv 0$: Assume we have found parameters satisfying (7.6)-(7.8). Since $\nu \equiv 0$, we can find the explicit solution of the system of equations (7.6):

$$\pi^* = \frac{\hat{\mu} - \hat{r}}{\sigma^2(1-\rho)},$$

and ρ is the smallest of the two roots of the quadratic

$$2\sigma^2(\hat{r}+\beta)\rho^2 - \big(2\sigma^2(\hat{r}+\beta) + (\hat{\mu}-\hat{r})^2 + 2(\delta+\beta\gamma)\sigma^2\big)\rho + 2(\delta+\beta\gamma)\sigma^2 = 0. \quad (7.11)$$

We see that ρ is the smallest root of the quadratic, because the coefficient of ρ^2 is positive, the left-hand side of (7.11) is negative when evaluated at 1, and we should have $\rho \leq 1$. Equation (7.11) can be solved for δ:

$$\delta = \frac{2\sigma^2(\hat{r}+\beta)\rho(1-\rho) + (\hat{\mu}-\hat{r})^2\rho - 2\beta\gamma\sigma^2(1-\rho)}{2\sigma^2(1-\rho)}. \quad (7.12)$$

We can also find an explicit expression for $k_0(\gamma)$:

$$\begin{aligned} k_0(\gamma) &= \max_{\pi\in[0,1]} \left[\gamma(\hat{r} + (\hat{\mu}-\hat{r})\pi) - \frac{1}{2}\sigma^2\pi^2\gamma(1-\gamma)\right] \\ &= \gamma\hat{r} + \frac{\gamma(\hat{\mu}-\hat{r})^2}{2\sigma^2(1-\gamma)}. \end{aligned} \quad (7.13)$$

We see that the optimal value of π in the expression for $k_0(\gamma)$ must lie in $(0,1)$, because $\rho > \gamma$ and $\pi^* \in (0,1)$. Using (7.12) and (7.13), we get

$$\rho(1-\gamma)\big(\delta - k_0(\gamma)\big) - (\rho-\gamma)(\delta+\beta\gamma) = -\gamma\rho(\rho-\gamma)(\hat{r}+\beta).$$

The right-hand side of this equation is negative, so

$$\rho(1-\gamma)\big(\delta - k_0(\gamma)\big) - (\rho-\gamma)(\delta+\beta\gamma) < 0,$$

and we see that (7.8) cannot hold. It follows that there are no parameter values that satisfy the given constraints. □

Now we will give an alternative proof of Lemma 7.1 for the case $\nu(dz) = \lambda\delta_{-\xi}(dz)$ and $\sigma = 0$. The solution of (7.6) for this case has been calculated in [8]. It is shown that if

$$\nu = \lambda\delta_{-\xi}(dz), \quad \xi \in (0,1), \quad \sigma = 0 \quad (7.14)$$

and

$$\mu > \hat{r}, \quad (\mu - \hat{r})e^{-(1-\rho)\xi} < \lambda(1-e^{-\xi}) < \mu - \hat{r}, \quad (7.15)$$

the constraint (7.6) can be written as

$$\begin{aligned} &\hat{\mu}-\hat{r}-\lambda(1-e^{-\xi})\Big(\big(1-\pi^*(1-e^{-\xi})\big)^{\rho-1}-1\Big)-\sigma^2(1-\gamma)\pi^*=0,\\ &\Big(\hat{r}+(\hat{\mu}-\hat{r})\pi^*+\beta\Big)\rho=\delta+\beta\gamma-\lambda\left((1-\pi^*(1-e^{-\xi}))^\rho-1+\rho\pi^*(1-e^{-\xi})\right), \end{aligned} \tag{7.16}$$

where $\hat{\mu}$ and π^* are given by

$$\hat{\mu}=\mu-\lambda(1-e^{-\xi})$$

and

$$\pi^*=\frac{1}{1-e^{-\xi}}\left(1-\left[\frac{\lambda(1-e^{-\xi})}{\mu-\hat{r}}\right]^{\frac{1}{1-\rho}}\right). \tag{7.17}$$

The following is an alternative proof of Lemma 7.1 for this case.

Alternative proof of Lemma 7.1 for $\nu(dz)=\lambda\delta_{-\xi}(dz)$: Assume there exist constants satisfying (7.8) and (7.14)-(7.17). We want to derive a contradiction by showing that $f>0$, where f is defined to be the difference between the right-hand side and the left-hand side of the second equation of (7.16), i.e.,

$$f(\varLambda)=\delta+\beta\gamma-\lambda\Big(\big(1-\pi^*(1-e^{-\xi})\big)^\rho-1+\rho\pi^*(1-e^{-\xi})\Big)-\big(\hat{r}+(\hat{\mu}-\hat{r})\pi^*+\beta\big)\rho,$$

where $\varLambda=(\delta,\beta,\gamma,\lambda,\rho,\mu,\pi^*,\xi,\hat{r})$. We will simplify the expression for f by introducing two new variables ζ and η:

$$\zeta=\left(\frac{\lambda\eta}{\mu-r}\right)^{\frac{1}{1-\rho}},\quad \eta=1-e^{-\xi}. \tag{7.18}$$

We have $\zeta\in(0,1)$ since $\pi^*=(1-\zeta)/\eta>0$, and $\eta\in(0,1)$ since $\xi\in(0,1)$. We introduce a new function $\tilde{f}$, which is equal to f, except that ζ and η are input arguments instead of π^*, $\hat{r}$ and ξ. Using (7.17) and (7.18), we see that

$$\tilde{f}(\tilde{\varLambda}):=f(\varLambda)=-\beta(\rho-\gamma)+\delta-\rho\lambda(1-\eta)\zeta^{\rho-1}-\lambda(1-\rho)\zeta^\rho+\lambda-\rho\mu,$$

where $\tilde{\varLambda}=(\delta,\beta,\gamma,\lambda,\rho,\mu,\zeta,\eta)$. Since $\hat{r}>0$ and $\zeta<1$, we see from the definition of ζ that

$$\mu>\lambda\eta\zeta^{\rho-1}. \tag{7.19}$$

From (7.15) we know that $(\mu-\hat{r})e^{-(1-\rho)\xi}<\lambda(1-e^{-\xi})$, and this implies that

$$\eta+\zeta>1. \tag{7.20}$$

Equation (7.8) implies that

$$\beta\le\frac{\delta(1-\rho)-\rho(1-\gamma)(\mu-\lambda\eta)}{\rho-\gamma}. \tag{7.21}$$

Since $\beta>0$, the numerator of this fraction must be positive, and this implies that

$$\delta>\frac{\rho(1-\gamma)}{1-\rho}(\mu-\lambda\eta). \tag{7.22}$$

We see that $\tilde{f}$ is decreasing in β, so using (7.21), we get

$$\tilde{f}(\tilde{\varLambda})\ge\delta\rho-\rho\zeta^{\rho-1}\lambda(1-\eta)-\zeta^\rho\lambda(1-\rho)-\lambda\rho\eta(1-\gamma)+\lambda-\rho\gamma\mu.$$

The right-hand side of this inequality is increasing in δ, so using (7.22), we get

$$\tilde{f}(\tilde{\Lambda}) > \frac{\mu\rho(\rho-\gamma) - \rho\lambda(1-\eta)(1-\rho)\zeta^{\rho-1} - \lambda(1-\rho)^2\zeta^\rho + \lambda(\rho\eta + \rho\gamma\eta - \rho + 1)}{1-\rho}.$$

The right-hand side of this inequality is increasing in μ, so using (7.19), we get

$$\tilde{f}(\tilde{\Lambda}) > \frac{\lambda}{1-\rho}\Big(\eta\rho(\zeta^{\rho-1}-1)(1-\gamma) - (1-\rho)^2\zeta^\rho + (1-\rho)(1-\rho\zeta^{\rho-1})\Big).$$

The right-hand side of this inequality is increasing in η, so (7.20) implies

$$\tilde{f}(\tilde{\Lambda}) > \frac{\lambda}{1-\rho}\Big(-\gamma\rho(1-\zeta)(\zeta^{\rho-1}-1) + \rho^2\zeta^{\rho-1} - (1-\rho+\rho^2)\zeta^\rho + \rho\zeta + 1 - 2\rho\Big).$$

The right-hand side of this inequality is decreasing in γ, so using $\gamma < \rho$, we get

$$\tilde{f}(\tilde{\Lambda}) > \lambda\Big(-\zeta^\rho + \zeta\rho + 1 - \rho\Big).$$

The right-hand side of this inequality is decreasing in ζ for $\zeta \in (0,1)$, and using $\zeta < 1$, we get

$$\tilde{f}(\tilde{\Lambda}) > 0.$$

We have obtained a contradiction, so there are no parameter values satisfying the given equations. □

Lemma 7.1 may suggest that (7.4) is not a valid solution formula, as there are no constants satisfying (7.6)-(7.8). However, looking more closely on the derivation of (7.8) in [8], we see that the constraint (7.8) may be weakened. The reason (7.8) was introduced, was to force

$$\widehat{F}\big(X, \widehat{V}, \widehat{V}_X, \widehat{V}_X^2, \mathscr{I}^\pi(X, \widehat{V})\big) \leq 0 \tag{7.23}$$

for $x \geq ky$. Replacing the constraint (7.8) by (7.23), we get a new and weaker system of constraints, which has solutions. We state in the lemma below that the constraint (7.23) is sufficient. Note that we may simplify (7.23) slightly by dividing by y^γ and replace the two variables x and y by the single variable $r := x/y$.

Lemma 7.2. *The value function $\widehat{V}$ defined by* (7.2) *is given by* (7.4)*, provided* (7.23) *holds and* (7.6) *has solutions satisfying* (7.7)*.*

If $\nu \equiv 0$ and (7.6) has solutions satisfying (7.7), (7.23) is automatically satisfied. The only constraint on the parameters in this case is therefore that (7.6) has a solution satisfying (7.7). We state and prove this result in the following lemma.

Lemma 7.3. *If $\nu \equiv 0$, the value function $\widehat{V}$ is given by* (7.4)*, provided* (7.6) *has a solution satisfying* (7.7)*.*

Proof. We need to prove that (7.23) is automatically satisfied if (7.6) has a solution satisfying (7.7). By inserting the expression (7.4) for $\widehat{V}$ in (7.23), multiplying by

$$\frac{(1+\beta k)^\gamma}{(y+\beta x)^\gamma} = \frac{1}{(y+\beta x)^\gamma}\frac{(1-\gamma)^\gamma}{(\rho-\gamma)^\gamma},$$

and using $\nu \equiv 0$ and the expression (7.5) for k, we see that (7.23) holds if and only if

$$f(\xi) \leq 0 \tag{7.24}$$

for all

$$\xi \in \left[\frac{\beta k}{1+\beta k}, 1\right] = \left[\frac{1-\rho}{1-\gamma}, 1\right],$$

where

$$\xi := \frac{\beta x}{y + \beta x} \tag{7.25}$$

and

$$\begin{aligned} f(\xi) :=& \frac{(1-\gamma)^\gamma}{(\rho-\gamma)^\gamma} \frac{(1-\xi)^\gamma}{\gamma} - \delta k_3 - \beta\gamma k_3(1-\xi) \\ &+ k_3 \max_{\pi\in[0,1]} \left[(\hat{r} + (\hat{\mu} - \hat{r})\pi)\gamma\xi - \frac{1}{2}\sigma^2\pi^2\xi^2\gamma(1-\gamma)\right]. \end{aligned} \tag{7.26}$$

We see from (7.25) that $x = ky$ is equivalent to $\xi = \frac{\beta k}{1+\beta k} = \frac{1-\rho}{1-\gamma}$. The optimal value of π in (7.26) is $\pi^* \in (0,1)$ for $x = ky$, and we see immediately from (7.26) that the optimal value of π decreases as ξ increases. Therefore we know that the optimal value of π in (7.26) is in $(0,1)$ for all $\xi \geq \frac{1-\rho}{1-\gamma}$, and π may be found by finding the maximum point of the parabola. We get

$$f(\xi) = \frac{(1-\gamma)^\gamma}{(\rho-\gamma)^\gamma} \frac{(1-\xi)^\gamma}{\gamma} - \delta k_3 - \beta\gamma k_3(1-\xi) + \hat{r}\gamma k_3 \xi + \frac{(\hat{\mu}-\hat{r})^2\gamma}{2\sigma^2(1-\gamma)}.$$

We have

$$\widehat{F}\big(X, \widehat{V}, \widehat{V}_X, \widehat{V}_X^2, \mathscr{I}^\pi(X, \widehat{V})\big) = 0$$

for $x = ky$ by (7.9), and therefore (7.24) is satisfied for $\xi = \frac{\beta k}{1+\beta k}$. We want to show that f is decreasing for $\xi \geq \frac{1-\rho}{1-\gamma}$. We have

$$\frac{\partial f}{\partial \xi}(\xi) = -\frac{(1-\gamma)^\gamma}{(\rho-\gamma)^\gamma}(1-\xi)^{\gamma-1} + k_3\gamma(\beta + \hat{r}),$$

which is decreasing in ξ. We have proved the lemma if we manage to show that

$$\frac{\partial f}{\partial \xi}\left(\frac{1-\rho}{1-\gamma}\right) \leq 0.$$

Inserting the expression (7.5) for k_3 and using (7.6) to find explicit expressions for δ and π^*, we get

$$\begin{aligned} \frac{\partial f}{\partial \xi}\left(\frac{1-\rho}{1-\gamma}\right) &= \frac{1-\gamma}{(\rho-\gamma)\big(\delta+\beta\gamma\big)}\big(-\delta + \beta\gamma + \rho(\beta+\hat{r})\big) \\ &= \frac{1-\gamma}{(\rho-\gamma)(\delta+\beta\gamma)}\rho\pi^*\left(-(\hat{\mu}-\hat{r}) + \frac{1}{2}\sigma^2\pi^*(1-\rho)\right) \\ &= -\frac{1-\gamma}{2(\rho-\gamma)(\delta+\beta\gamma)}\rho\pi^*(\hat{\mu}-\hat{r}) \\ &\leq 0, \end{aligned}$$

and the lemma is proved. □

If $\nu(dz) = \lambda\delta_{-\xi}(dz)$, the system of constraints (7.6), (7.7) and (7.23) also has solutions, for example the set of parameters we will consider in Chapter 15.

Now we will use the explicit solution formula in [8] to find an explicit solution formula also in the finite-horizon case. It turns out that the value function of our problem, defined by (3.4), is given by

$$V(t,x,y) = e^{-\delta t}\widehat{V}(x,y) = \begin{cases} e^{-\delta t}\left(k_1 y^\gamma + k_2 y^\gamma \left[\frac{x}{ky}\right]^\rho\right) & \text{if } 0 \le x < ky, \\ k_3 e^{-\delta t}\left(\frac{y+\beta x}{1+\beta k}\right)^\gamma & \text{if } x \ge ky \ge 0, \end{cases} \tag{7.27}$$

if U is on the form (7.1), and the terminal utility function is defined by

$$W(x,y) = \begin{cases} e^{-\delta T}\left(k_1 y^\gamma + k_2 y^\gamma \left[\frac{x}{k}y\right]^\rho\right) & \text{if } 0 \le x < ky, \\ k_3 e^{-\delta T}\left(\frac{y+\beta x}{1+\beta k}\right)^\gamma & \text{if } x \ge ky \ge 0, \end{cases} \tag{7.28}$$

where the constants k_1, k_2, k_3, k_4, ρ and π^* satisfy (7.5) and (7.6). Again the only constraints the parameters must satisfy are (7.7) and (7.23).

Theorem 7.4. *If U is defined by* (7.1)*, W is defined by* (7.28)*,* (7.23) *holds and* (7.6) *has a solution satisfying* (7.7)*, the value function V is given by* (7.27)*. The optimal consumption strategy is to keep a constant fraction π^* of the wealth invested in the risky asset, where π^* is defined by the system of equations* (7.6)*.*

Proof. Define $v : \overline{\mathscr{D}_T}$ by

$$v(t,x,y) := \begin{cases} e^{-\delta t}\left(k_1 y^\gamma + k_2 y^\gamma \left[\frac{x}{ky}\right]^\rho\right) & \text{if } 0 \le x < ky, \\ k_3 e^{-\delta t}\left(\frac{y+\beta x}{1+\beta k}\right)^\gamma & \text{if } x \ge ky \ge 0. \end{cases}$$

We want to show that $v = V$. The first step is to prove that v is a classical solution of (5.11). We see that $v \in C^{1,2,1}(\overline{\mathscr{D}_T})$ by direct computations. We have

$$G(D_X v) = e^{-\delta t}\widehat{G}(D_X \widehat{V}) \tag{7.29}$$

and

$$\begin{aligned} &v_t + F\big(t, X, D_X v, D_X^2 v, \mathscr{I}^\pi(t,X,v)\big) \\ &\quad = e^{-\delta t}\Big(\widehat{V}_t + \widehat{F}\big(X, D_X\widehat{V}, D_X^2\widehat{V}, \mathscr{I}^\pi(X,\widehat{V})\big)\Big). \end{aligned} \tag{7.30}$$

for all $(t,X) \in \overline{\mathscr{D}_T}$. In [8] the authors show that (7.4) is a classical solution of (7.3). Using (7.29) and (7.30), we see that this is equivalent to (7.27) being a classical solution of (5.11).

We see by insertion that v satisfies the terminal condition

$$v(T,x,y) = \begin{cases} e^{-\delta T}\left(k_1 y^\gamma + k_2 y^\gamma \left[\frac{x}{ky}\right]^\rho\right) & \text{if } 0 \le x < ky, \\ k_3 e^{-\delta T}\left(\frac{y+\beta x}{1+\beta k}\right)^\gamma & \text{if } x \ge ky \ge 0, \end{cases}$$

so $v(T,x,y) = W(x,y)$ for all $(x,y) \in \overline{\mathscr{D}}$. We know that V satisfies (4.2), so we need to prove that

$$W(x,y) = \max_{c\in[0,x]} W(x-c, y+\beta c) \tag{7.31}$$

for all $(x,y) \in \overline{\mathscr{D}}$. For $x < ky$ we have

$$\frac{dW}{dc}(x-c, y+\beta c) = \frac{(y+\beta c)^{\gamma-1}\beta}{\delta+\beta\gamma} g(z),$$

where $g : [0, \frac{x}{ky}] \to \mathbb{R}$ is defined by

$$g(z) = 1 - (1-\rho)z^{\rho} - \rho z^{\rho-1},$$

and $z = \frac{x-c}{k(y+\beta c)}$. We see that $\frac{dW}{dc}(x-c, y+\beta c)$ has the same sign as $g(z)$. We see easily that $g(z) \leq 0$ on $\left[0, \frac{x}{ky}\right]$, and therefore $W(x,y) \geq W(x-c, y+\beta c)$ for $c \in [0,x]$, and (7.31) is satisfied.

For $x \geq ky$, $W(x,y) = W(x-c, y+\beta c)$ as long as $(x-c) > k(y+\beta c)$, and $W(x,y) \geq W(x-c, t+\beta c)$ for $(x-c) \leq k(y+\beta c)$. Therefore (7.31) is satisfied also in this case.

By Lemma 4.9, we see that v is a viscosity solution of (5.11). Using Theorems 6.14 and 6.16, we see that $V = v$. □

Note that W could be replaced by any other function $\tilde{W}$ satisfying

$$W(x,y) = \max_{c\in[0,x]} \tilde{W}(x-c, y+\beta c).$$

An optimal consumption process is given by almost exactly the same formula as for the infinite-horizon problem.

Theorem 7.5. *An optimal consumption process C_s^* is given by*

$$C_s^* = \Delta C_t^* + \int_t^s \frac{X_{s'}^{\pi^*,C^*}}{1+\beta k}\, dZ_{s'},$$

where

$$\Delta C_t^* = \left[\frac{x - kY_{t-}}{1+\beta k}\right]^+, \tag{7.32}$$

$$Z_s = \sup_{t\leq s'\leq s}\left[\ln \frac{X_{s'}^{\pi^*,\Delta C_t^0}}{Y_{s'}^{\pi^*,\Delta C_t^0}} - \ln k\right]^+,$$

$$Y_s^{\pi^*,\Delta C_t^0} = (Y_t + \beta \Delta C_t^*)e^{-\beta s}$$

and

$$X_s^{\pi^*,\Delta C_t^0} = (x - \Delta C_t^*) + \int_t^s (\hat{r} + (\hat{\mu}-\hat{r})\pi^*)X_{s'}^{\pi^*,\Delta C_t^0}\, ds' + \int_t^s \sigma\pi^* X_{s'}^{\pi^*,\Delta C_t^0}\, dB_{s'}$$
$$+ \int_t^s \pi^* X_{s'-}^{\pi^*,\Delta C_t^0} \int_{-\infty}^{0-} (e^z - 1)\, \tilde{N}(ds', dz).$$

Proof. First we will show that a so-called k-ratio strategy is optimal. A k-ratio strategy is a strategy where the control keeps the ratio $X_s^{\pi,C}/Y_s^{\pi,C}$ less than or equal to some constant k for all $s \in [t,T]$ almost certainly. If $X^{\pi,C}/Y^{\pi,C} > k$, the agent will consume such that $X^{\pi,C}/Y^{\pi,C} = k$, and if $X^{\pi,C}/Y^{\pi,C} < k$, no consumption will occur.

From the explicit formula for V, we see that $G(D_X V) = 0$ when $X_s^{\pi,C} \geq kY_s^{\pi,C}$. This implies that $V(X_s^{\pi,C}, Y_s^{\pi,C}) = V(X_s^{\pi,C} - c, Y_s^{\pi,C} + \beta c)$ for $X_s^{\pi,C} \geq kY_s^{\pi,C}$ for all $c \in [0, X_s^{\pi,C}]$ satisfying $(X_s^{\pi,C} - c) \geq k(Y_s^{\pi,C} + \beta c)$. We see from Lemma 4.8 that it is optimal to have consumption when $X_s^{\pi,C} > kY_s^{\pi,C}$.

When $X^{\pi,C}/Y^{\pi,C} > k$, on the other hand, we see from the explicit formula for V that $G(D_X V) < 0$. The inequality $G(D_X V) < 0$ is equivalent to $V(X_s^{\pi,C}, Y_s^{\pi,C}) > V(X_s^{\pi,C} - c, Y_s^{\pi,C} + \beta c)$ for all $c \in (0, X_s^{\pi,C}]$, and we see from Lemma 4.8 that it is optimal not to consume. It follows that a k-ratio strategy is optimal.

We get an initial consumption gulp if $x > ky$. After the consumption has been performed, we should have $X_t^{\pi,C} = kY_t^{\pi,C}$, and we see that the size of the consumption gulp ΔC_t^0 is given by (7.32).

A consumption plan is optimal if it satisfies (i) $X_s^{\pi,C}/Y_s^{\pi,C} \leq k$ almost certainly for all $s \in [t,T]$, and (ii) there is no consumption when $X_s^{\pi,C}/Y_s^{\pi,C} < k$. The constraint (i) is equivalent to

$$\ln \frac{X_s^{\pi^*,C^*}}{Y_s^{\pi^*,C^*}} \leq \ln k.$$

Define

$$Z_s = \sup_{0 \leq s' \leq s} \left[\ln \frac{\widehat{X}_{s'}^{\pi^*,\Delta C_t^0}}{\widehat{Y}_{s'}^{\pi^*,\Delta C_t^0}} - \ln k \right]^+ . \tag{7.33}$$

Note that $X_s^{\pi^*,\Delta C_t^0}$ and $Y_s^{\pi^*,\Delta C_t^0}$ are unregulated processes, i.e., they take the value $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ would have taken if no consumption unless the initial consumption gulp was performed. The constraint (i) is satisfied if the process $\ln(X_s^*/Y_s^*)$ satisfies

$$\ln \frac{X_s^{\pi^*,C^*}}{Y_s^{\pi^*,C^*}} = \ln \frac{X_s^{\pi^*,\Delta C_t^0}}{Y_s^{\pi^*,\Delta C_t^0}} - Z_s, \tag{7.34}$$

because

$$\ln \frac{X_s^{\pi^*,\Delta C_t^0}}{Y_s^{\pi^*,\Delta C_t^0}} - Z_s \leq \frac{X_s^{\pi^*,\Delta C_t^0}}{Y_s^{\pi^*,\Delta C_t^0}} - \left[\frac{X_s^{\pi^*,\Delta C_t^0}}{Y_s^{\pi^*,\Delta C_t^0}} - \ln k \right] = \ln k.$$

We want to prove that there is a strategy C^* such that (7.34) and (ii) are satisfied.

First note that Z is non-decreasing, and increases only when $X_s^* = kY_s^*$ if (7.34) is correct. By a differential form of Itô's formula (Theorem (2.22)), we have

$$\begin{aligned} d \ln \frac{X_s^{\pi^*,C^*}}{Y_s^{\pi^*,C^*}} = & \left(\hat{r} + \beta(\hat{\mu} - \hat{r})\pi^* - \frac{1}{2}(\sigma\pi^*)^2 \right) ds - \left(\frac{1}{X_s^{\pi^*,C^*}} + \frac{\beta}{Y_s^{\pi^*,C^*}} \right) dC_s^* \\ & + \int_{-\infty}^{0-} \ln\big(1 + \pi^*(e^z - 1)\big)\, \tilde{N}(ds, dz) + \sigma\pi^*\, dB_s \\ & + \int_{-\infty}^{0-} \Big(\ln\big(1 + \pi^*(e^z - 1)\big) - \pi^*(e^z - 1) \Big)\, \nu(dz) \end{aligned}$$

and

$$d\ln\frac{X_s^{\pi^*,\Delta C_t^*}}{Y_s^{\pi^*,\Delta C_t^*}} = \Big(\hat{r}+\beta(\hat{\mu}-\hat{r})\pi^* - \frac{1}{2}(\sigma\pi^*)^2\Big)\,ds - \left(\frac{1}{X_s^{\pi^*,\Delta C_t^*}}+\frac{\beta}{Y_s^{\pi^*,\Delta C_t^*}}\right)dC_s^*$$
$$+\int_{-\infty}^{0^-}\ln\big(1+\pi^*(e^z-1)\big)\,\tilde{N}(ds,dz)+\sigma\pi^*\,dB_s$$
$$+\int_{-\infty}^{0^-}\Big(\ln\big(1+\pi^*(e^z-1)\big)-\pi^*(e^z-1)\Big)\,\nu(dz).$$

We see that (7.34) is satisfied if

$$dZ_s = \left(\frac{1}{X_{s'}^{\pi^*,C^*}}+\frac{\beta}{Y_{s'}^{\pi^*,C^*}}\right)dC_s^*,$$

i.e.,

$$Z_s = \int_0^t \frac{Y_{s'}^{\pi^*,C^*}+\beta X_{s'}^{\pi^*,C^*}}{X_{s'}^{\pi^*,C^*}Y_{s'}^{\pi^*,C^*}}\,dC_{s'}^*.$$

This is equivalent to

$$C_s^* = \int_0^s \frac{X_{s'}^{\pi^*,C^*}Y_{s'}^{\pi^*,C^*}}{Y_{s'}^{\pi^*,C^*}+\beta X_{s'}^{\pi^*,C^*}}\,dZ_{s'} = \int_0^s \frac{X_{s'}^{\pi^*,C^*}}{1+\beta k}\,dZ_{s'},$$

where the last equality follows because Z_s only increases when $X_s^{\pi^*,C^*}/Y_s^{\pi^*,C^*}=k$. We see that C^* satisfies (ii), because C^* only increases when Z increases. We also see that $(\pi^*,C^*)\in\mathscr{A}_{t,x,y}$, since C^* is increasing, and we have proved that C^* is an optimal consumption strategy. □

We see from Theorems 7.4 and 7.5 that both the consumption strategy and the distribution of wealth between the risky and safe asset, are independent of the time t.

We try searching for explicit solution formulas for other functions W by two different strategies:

1. Assume the solution has the form (7.27), but where k_1, k_2, k_3, k_4 and k are assumed to be functions of t instead of constants.
2. Assume π is constant, k is a function of t, and that the solution is separable on each of the domains $x>ky$ and $x<ky$.

In both cases we end up with same solution formula as before. Note that the second strategy includes solutions on a form similar to the one we find for the Merton problem below.

7.2 The Merton problem

In this section we consider the Merton problem with consumption, where the stock price is modelled as a geometric Lévy process. The Merton problem can be thought of as the limiting case when $\beta\to\infty$ in the problem described in Chapter 3, i.e., we want to optimize utility from *present* consumption, instead of optimizing average past consumption. The consumption process is assumed to be absolutely continuous, and can therefore be written on the form $C_s = \int_t^s c_{s'}\, ds'$, where $c_{s'}$ is the consumption rate at time s'.

An explicit solution formula for the infinite-horizon version of the problem is described in [8]. A solution formula for the infinite-horizon problem is also found in [26] and, with transaction costs, [27], but with a slightly more restrictive condition on the Lévy measure in a neighbourhood of zero. No articles containing an explicit solution formulas for the finite-horizon case was found in the research work for this thesis.

The problem solved in [8] can be stated as

$$\widehat{V}^M(x) = \sup_{(c,\pi)\in\mathscr{A}_x} \mathrm{E}\left[\int_0^\infty e^{-\delta t}\frac{c_t^\gamma}{\gamma}dt\right],$$

for $x \geq 0$, where the wealth process is given by

$$dX_t = (\hat{r} + (\hat{\mu} - \hat{r})\pi_t)X_t dt - c_t dt + \sigma X_t \pi_t dB_t + X_{t^-}\pi_{t^-}\int_{\mathbb{R}\setminus\{0\}}(e^z - 1)\tilde{N}(dt, dz)$$

with initial value $X_0 = x$, ν satisfies (3.1), and the set of admissible controls $\mathscr{A}_x$ is defined as in the finite-horizon case presented below. The HJB equation for this problem is

$$\begin{aligned} 0 = \max_{c\geq 0,\, \pi\in[0,1]} \Bigg[&\frac{c^\gamma}{\gamma} - \delta\widehat{V}^M + \big(\hat{r} + \pi(\hat{\mu} - \hat{r})\big)x\widehat{V}_x^M - c\widehat{V}_x + \frac{1}{2}(\sigma\pi x)^2\widehat{V}_{xx}^M \\ &+ \int_{\mathbb{R}\setminus\{0\}} \widehat{V}^M\big(t, x + \pi x(e^z - 1)\big) - \widehat{V}^M(t,x) - \pi x(e^z - 1)\widehat{V}_x^M(t,x)\,\nu(dz)\Bigg]. \end{aligned}$$

It is proved in [8] that

$$\widehat{V}^M(x) = Kx^\gamma,$$

where

$$K = \frac{1}{\gamma}\left[\frac{1-\gamma}{\delta - k(\gamma)}\right]^{1-\gamma}. \tag{7.35}$$

It is also proved that the optimal consumption rate c is given by

$$c = (K\gamma)^{\frac{1}{\gamma-1}}x,$$

and that it is optimal to let π be constantly equal to the solution of this equation:

$$(\hat{\mu} - \hat{r}) - (1-\gamma)\sigma^2\pi + \int_{\mathbb{R}\setminus\{0\}}\Big(\big(1 + \pi(e^z - 1)\big)^{\gamma-1} - 1\Big)(e^z - 1)\,\nu(dz) = 0.$$

The only conditions on the parameters, in addition to those given in Chapter 3, is that $\delta > k(\gamma)$ and

$$\int_{\mathbb{R}\setminus\{0\}}\left(1 - e^{-(1-\gamma)z}\right)(e^z - 1)\,\nu(dz) > (\hat{\mu} - \hat{r}) - (1-\gamma)\sigma^2. \tag{7.36}$$

We will consider a finite-horizon version of the above problem, where the value function is defined by:

$$V^M(t,x) := \sup_{(c,\pi)\in\mathscr{A}_x} \mathrm{E}\left[\int_t^T e^{-\delta s}\frac{c_s^\gamma}{\gamma}\,ds + a\frac{1}{\gamma}\left(X_T^{\pi,C}\right)^\gamma\right], \tag{7.37}$$

where $(t,x)\in[0,T]\times[0,\infty)$, $a>0$ is a constant, the wealth process is given by

$$dX_s = (\hat{r}+(\hat{\mu}-\hat{r})\pi_s)X_s\,ds - c_s\,ds + \sigma X_s\pi_s dB_s + X_{s^-}\pi_{s^-}\int_{\mathbb{R}\backslash\{0\}}(e^z-1)\,\tilde{N}(ds,dz)$$

with initial value $X_t = x$, and the set of admissible controls $\mathscr{A}_{t,x}$ is defined as follows: The control $(\pi,c)\in\mathscr{A}_{t,x}$ if

(cm_i) c_s is a positive and adapted process on $[t,T]$ such that $\mathrm{E}\big[\int_t^s c_{s'}\,ds'\big]<\infty$ for all $s\in[t,T]$,

(cm_{ii}) π_s is an adapted càdlàg process on $[t,T]$ with values in $[0,1]$, and

(cm_{iii}) c_s is such that $X_s^{\pi,c}\geq 0$ almost everywhere for all $s\in[t,T]$.

The HJB equation for this problem is

$$\begin{aligned}0 = \max_{c\geq 0,\,\pi\in[0,1]}\Bigg[&e^{-\delta t}\frac{c^\gamma}{\gamma} + V_t^M + \big(\hat{r}+\pi(\hat{\mu}-\hat{r})\big)xV_x^M - cV_x^M + \frac{1}{2}(\sigma\pi x)^2V_{xx}^M\\ &+\int_{\mathbb{R}\backslash\{0\}} V^M\big(t,x+\pi x(e^z-1)\big) - V^M(t,x) - \pi x(e^z-1)V_x^M(t,x)\,\nu(dz)\Bigg],\end{aligned} \tag{7.38}$$

for all $(t,x)\in[0,T]\times(0,\infty)$, and the terminal condition is

$$V^M(T,x) = a\frac{x^\gamma}{\gamma} \tag{7.39}$$

for all $x>0$.

Inspired by the result in section 7.1, we first try to find a solution on the form $e^{-\delta t}\widehat{V}^M$, where $\widehat{V}^M$ solves the infinite-horizon problem. We see by insertion in (7.38) that $e^{-\delta t}\widehat{V}^M$ is a solution, and see from (7.39) that the constant a is given by

$$a = \left[\frac{1-\gamma}{\delta-k(\gamma)}\right]^{1-\gamma}.$$

It follows that $V^M = e^{-\delta t}\widehat{V}$ if a satisfies this equation.

Now we try to find solutions of (7.38) and (7.39) for other values of a. We try to find a solution on the form

$$V^M(t,x) = \big(A+Be^{D(T-t)}\big)^{1-\gamma}x^\gamma e^{-\delta t}, \tag{7.40}$$

for constants $A,B,D\in\mathbb{R}$. This form is inspired by the classical article [37] of Merton, where it is shown that the solution of the finite-horizon problem

$$\tilde{V}^M(t,x) =: \sup_{(\pi,c)\in\mathscr{A}_{x,t}} \mathrm{E}\left[\int_t^T e^{-\delta s}\frac{(c_s)^\gamma}{\gamma}\,ds\right]$$

in the case $\nu\equiv 0$, is given by

$$\tilde{V}^M(x,t) = \frac{1}{\gamma}e^{-\delta t}\left[\frac{1-\gamma}{\delta-\gamma\alpha}e^{-(\delta-\gamma\alpha)(T-t)/\delta}\right]^{1-\gamma}x^\gamma, \tag{7.41}$$

where $\alpha = \hat{r} + (\hat{\mu} - \hat{r})^2/(2\sigma^2(1-\gamma))$[1].

Differentiating the argument of the max operator in (7.38) with respect to c, and then inserting the expression (7.40) for V^M, we see that the optimal consumption rate is

$$c^* = (e^{\delta t} V_x)^{\frac{1}{\gamma-1}} = \gamma^{\frac{1}{\gamma-1}} \frac{x}{A + Be^{D(T-t)}}. \tag{7.42}$$

We now assume the optimal value $\pi^* = \pi^*(t,x)$ of π is in the interval $(0,1)$. Differentiating the argument of the max operator in (7.38) with respect to π, and then inserting the expression (7.40) for V^M, we get

$$(\hat{\mu} - \hat{r}) - \sigma^2(1-\gamma)\pi^* + \int_{\mathbb{R}\backslash\{0\}} \Big(\big(1 + \pi^*(e^z - 1)\big)^\gamma - 1 \Big)(e^z - 1)\, \nu(dz) = 0. \tag{7.43}$$

We note that π^* is independent of x and t. It is shown in [8] that (7.43) has a unique solution in $(0,1)$ if (7.36) holds. Inserting the expression for c^* into (7.38), we obtain the following equation

$$\Big\{ \gamma^{-\frac{1}{1-\gamma}} - \gamma^{-\frac{\gamma}{1-\gamma}} - A\big(\delta - k(\gamma)\big) \Big\} - Be^{D(T-t)} \Big\{ D(1-\gamma) + (\delta - k(\gamma)) \Big\} = 0.$$

The expression on the left-hand side must be 0 for all values of t. Therefore we see that

$$A = \frac{\gamma^{\frac{-1}{1-\gamma}}(1-\gamma)}{\delta - k(\gamma)} \tag{7.44}$$

and

$$D = -\frac{\delta - k(\gamma)}{1-\gamma}. \tag{7.45}$$

We see immediately from (7.44) and (7.45) that the parameters of the problem must satisfy

$$\delta \neq k(\gamma), \tag{7.46}$$

and that $AD = -\gamma^{-\frac{1}{1-\gamma}}$. Note that the constraint $\delta > k(\gamma)$, which must hold for the infinite-horizon problem in order for K to be defined, is not necessary in this case. The constraint $\delta > k(\gamma)$ may be necessary to find estimates on the growth of V^M, and hence to prove that the integral of the HJB equation is well-defined, but the constraint $\delta > k(\gamma)$ is not necessary for all Lévy measures. We can find an expression for B by using the terminal value condition (7.39):

$$(A + B)^{1-\gamma} = a/\gamma. \tag{7.47}$$

We need to check that (7.40) is well-defined for all t, i.e.,

$$A + Be^{D(T-t)} \geq 0 \tag{7.48}$$

for all $t \in [0,T]$. By (7.39), we know that (7.48) is satisfied for $t = T$. We see directly from (7.44) and (7.45) that A and D have opposite signs. If $D < 0$, $|Be^{D(T-t)}|$ is increasing in t, and therefore $A + Be^{D(T-t)}$ must stay positive when t decreases from T. If $D > 0$, $A < 0$, and we see from (7.47) that $B > 0$. This implies that $Be^{D(T-t)}$ is decreasing in t, and therefore $A + Be^{D(T-t)} \geq 0$ for all values of $t \in [0,T]$.

We can summarize the results of this section in the following theorem.

[1] There was a printing error in [37], and therefore the exact formulation of equation (7.41) in this thesis is from [42] instead of [37].

Theorem 7.6. *The value function defined by* (7.37) *is given by* (7.40) *provided* (7.36) *and* (7.46) *holds. The constants A, B and D are defined by* (7.44)*,* (7.47) *and* (7.45)*, respectively. The optimal control strategy is given by* (7.42) *and* (7.43)*.*

We note that V^M converges to $e^{-\delta t}\widehat{V}^M$ as $T \to \infty$ if $\delta > k(\gamma)$: If $\delta > k(\gamma)$, we see from (7.45) that $D < 0$, and it follows from (7.40) that

$$\lim_{T\to t} \hat{V}^M(t,x) = A^{1-\gamma}x^\gamma e^{-\delta t} = e^{-\delta t}\widehat{V}^M(t,x).$$

This corresponds well with a financial interpretation of the problem, as the terminal utility function becomes less important as the time-horizon increases.

It is shown in [8] that $\pi^* < \pi^*_{GBM}$, where

$$\pi^*_{GBM} = \frac{\mu + \sigma^2/2 - \hat{r}}{(1-\gamma)\sigma^2}$$

is the optimal value of π for $\nu \equiv 0$. The case $\nu \equiv 0$ is the case where the Lévy process is replaced by a Brownian motion. That $\pi^* < \pi^*_{GBM}$ is reasonable from a financial interpretation of the problem: An increasing, concave utility function gives priority to investments with low risks, and to investments with high expected return. Processes with jumps add more uncertainty to the problem, so the investor will invest more in the safe asset, at the expense of investing less in the asset with highest expected return.

The fraction of wealth that should be consumed per unit time, c^*/x, is an increasing function of time if $D > 0$. We see this since $D > 0$ implies $A < 0$ and $B > 0$, so the term $Be^{D(T-t)}$ is decreasing. If $D < 0$, on the other hand, c^*/x may be both an increasing and a decreasing function of time: We have $B > 0$ if a is sufficiently large, so c^*/x becomes a decreasing function of time. If a is sufficiently small, we get $B < 0$, so c^*/x is increasing.

We can explain this behaviour from a financial point of view: In many cases it is advantageous to increase the ratio c^*/x with time, because the wealth is expected to increase in the absence of consumption. By delaying the consumption, we can probably obtain a larger total consumption. If δ and a are large, however, the agent gives priority to early consumption and large final wealth. In this case, it may be advantageous to decrease the rate c^*/x with time.

In Section 4 we saw that V is not generally increasing or decreasing in t. The problem considered in this section is not exactly the same as the problem considered in the rest of the thesis, but the solution found here can still illustrate how V may vary with t for general stochastic control problems. Generally we expect V^M to decrease with time for fixed x, as $X_s^{\pi,C}$ is expected to increase with time in the absence of consumption. However, as we will see in the paragraph below, V_t^M may be positive for certain parameter values.

For $D > 0$ we see immediately from (7.40) that V^M is decreasing in t. We get the same result in the case $D < 0$ and $B < 0$. In the case $D < 0$ and $B > 0$, however, $V_t > 0$ if B is sufficiently large. We see by a calculation of V_t^M that there is a constant t^*, such that $V_t^M > 0$ for $t < t^*$ and $V_t^M < 0$ for $t > t^*$. We may have $t^* > T$, such that $V_t^M > 0$ for all $t \in [0,T]$.

We can explain the occurrence of $V_t^M > 0$ from a financial point as follows: We may have $V_t^M > 0$ if $D < 0$ and $B > 0$, i.e., if δ, σ and/or a are large. If δ and a are large, the agent gets much satisfaction from final wealth compared to satisfaction from consumption. A large value of σ tells us that the uncertainties in the marked are large, and the uncertainty of the final wealth grows as we go backwards in time. The agent gives priority to safe investments, and therefore it may be better to have a guaranteed wealth of x at a later time, instead of having wealth x at an early time.

Part II
A penalty approximation to the optimization problem

The HJB equation derived in Chapter 5 is difficult to solve, because it is non-linear, non-local and is subject to a gradient constraint. It is especially the gradient constraint G that makes the equation difficult to solve numerically. A problem containing only an HJB operator similar to F, is easier to solve than a problem containing the maximum between two functions.

The gradient constraint of the problem arises because the control problem is singular, i.e., because we allow the cumulative consumption function C to be discontinuous. We will consider a continuous version of the problem, where the control C is absolutely continuous and has a derivative that is bounded by $1/\epsilon$. This problem corresponds to an HJB equation $(v_\epsilon)_t + F_\epsilon = 0$, where $F_\epsilon = F + \frac{1}{\epsilon}\max\{G; 0\}$. We will prove that $(v_\epsilon)_t + F_\epsilon = 0$ has a unique viscosity solution V_ϵ, and that V_ϵ converges uniformly to V as $\epsilon \to 0$.

The theory developed in this part will be used in Part IV when constructing a numerical scheme for solving the HJB equation. Instead of solving the original HJB equation, we will solve the penalty approximation $(v_\epsilon)_t + F_\epsilon = 0$ for a small value of ϵ. The convergence results proved in this part will guarantee that the numerical solution converges to V as $\epsilon \to 0$.

In Chapter 8 the penalty approximation $(v_\epsilon)_t + F_\epsilon = 0$ is derived. In Chapter 9 viscosity theory for the penalty problem is developed, and some properties of the family of functions $\{V_\epsilon\}_{\epsilon>0}$ is proved. In Chapter 10 it is proved that V_ϵ converges uniformly to V on bounded subsets of $[0,T) \times \overline{\mathscr{D}}$.

Chapter 8
Deriving a penalty approximation

In this section we consider a non-singular version of the original control problem. We derive a penalty approximation to the original HJB equation by finding the HJB equation of the non-singular problem. The original problem is singular because the displacement of the state variables $X^{\pi,C}$ and $Y^{\pi,C}$ due to control effort may be discontinuous, and in this chapter we will assume the control C is absolutely continuous, i.e., the displacement of $X^{\pi,C}$ and $Y^{\pi,C}$ due to the control C is differentiable in time almost everywhere.

The technique is inspired by [17], where a singular portfolio optimization problem with transaction costs is considered. A penalty approximation to the HJB equation is constructed by considering a non-singular version of the original problem, and the result is a penalty approximation on a similar form to the one we will derive below.

For any $\epsilon > 0$, suppose $C_s = \int_t^s c(s')\, ds'$ for all $s \in [t,T]$, where $|c(s')| \leq 1/\epsilon$ almost everywhere, i.e., C is absolutely continuous with bounded derivative. Define

$$V_\epsilon(t,x,y) = \sup_{(\pi,C)\in\mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right], \tag{8.1}$$

where

$$\begin{aligned}\mathscr{B}^\epsilon_{t,x,y} = \{&(\pi,C) \in \mathscr{A}_{t,x,y} :\ C \text{ is absolutely continuous} \\ &\text{with derivative bounded by } 1/\epsilon \text{ a.e.}\}.\end{aligned} \tag{8.2}$$

We follow the steps of the derivation in Chapter 5. We insert equation (5.5) and $dC_s = c(s)ds$ into equation (5.4), and let Δt go towards 0, to get

$$\begin{aligned}0 = {}& e^{-\delta s}U(y) + (V_\epsilon)_t + \sup_{\pi\in[0,1]}\left[\big(\hat{r}(1-\pi) + \pi\hat{\mu}\big)x(V_\epsilon)_x + \frac{1}{2}(\sigma\pi x)^2(V_\epsilon)_{xx} + \mathscr{I}^\pi(t,X,V_\epsilon)\right] \\ &+ \sup_{c\in[0,1/\epsilon]}\big[c(-(V_\epsilon)_x + \beta(V_\epsilon)_y)\big].\end{aligned}$$

This equation can be solved explicitly for c: If $-V_x + \beta V_y \geq 0$, the maximum is obtained for $c = 1/\epsilon$, and if $-V_x + \beta V_y \leq 0$, the maximum is obtained for $c = 0$. Inserting this value for c, we obtain the following HJB equation:

$$(v_\epsilon)_t + F_\epsilon\big(t, X, D_X v_\epsilon, D_X^2 v_\epsilon, \mathscr{I}^\pi(t,X,v_\epsilon)\big) = 0, \tag{8.3}$$

where

$$F_\epsilon\big(t, X, D_X v_\epsilon, D_X^2 v_\epsilon, \mathscr{I}^\pi(t, X, v_\epsilon)\big) = F\big(t, X, D_X v_\epsilon, D_X^2 v_\epsilon, \mathscr{I}^\pi(t, X, v_\epsilon)\big) + \frac{1}{\epsilon}\max\{G(D_X v_\epsilon);\, 0\},$$

and F and G are defined by (5.12) and (5.13), respectively. Equation (8.3) is the penalty approximation of (5.11) that we will use in this thesis.

The terminal condition becomes

$$V_\epsilon(T, x, y) = W(x, y) \tag{8.4}$$

by (8.1), because the consumption is continuous. If we replace the terminal utility function W by $\widehat{W}$ as described in Section 4.1, we obtain the following terminal condition:

$$V_\epsilon(T, x, y) = \widehat{W}(x, y). \tag{8.5}$$

Note that the terminal condition is similar for the original problem and the penalty problem if (8.5) holds, while it may be different otherwise. If the terminal utility function is defined by (8.5), V_ϵ converges to V on the whole domain $\overline{\mathscr{D}_T}$, but if the terminal utility function is given by (8.4), V_ϵ only converges to V on $[0, T) \times \overline{\mathscr{D}}$. We will always assume the terminal condition is (8.4) unless otherwise stated. Viscosity solution theory can be identically developed for the two terminal conditions, but the equicontinuity results we will develop in Section 9.2, are different. In Section 9.2 we develop results for the terminal condition (8.4), and the corresponding results for (8.5) will not be developed before in Section 10.4.

For $x = 0$ the boundary condition (4.1) is still valid.

Just as for the singular problem, we will assume the dynamic programming principle and the existence of optimal controls. Lemma 8.1 corresponds to Lemma 3.5 in the singular case, and references to proofs of the dynamic programming principle are given after Lemma 3.5.

Lemma 8.1 (The Dynamic Programming Principle). *For all $(t, x, y) \in \overline{\mathscr{D}_T}$, $\epsilon > 0$ and $\Delta t \in [0, T - t]$, we have*

$$V_\epsilon(t, x, y) = \sup_{(\pi, C) \in \mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_t^{t+\Delta t} e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + V_\epsilon\left(t + \Delta t, X_{t+\Delta t}^{\pi,C}, Y_{y+\Delta}^{\pi,C}\right)\right].$$

Lemma 8.2 corresponds to Lemma 3.6 in the singular case. A proof of the existence of optimal controls for a non-singular control problem, is given in Section I.11.1 in the second edition of [25].

Lemma 8.2 (Existence of Optimal Controls). *For all $(t, x, y) \in \overline{\mathscr{D}_T}$ and $\epsilon > 0$, there exists a $(\pi^*, C^*) \in \mathscr{B}^\epsilon_{t,x,y}$ such that*

$$V_\epsilon(t, x, y) = \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi^*,C^*}\right) ds + W\left(X_T^{\pi^*,C^*}, Y_T^{\pi^*,C^*}\right)\right].$$

Chapter 9
Properties of the penalty approximation

In this section we will prove some properties of the value function V_ϵ of the penalized problem. We have divided the properties into two classes: First we consider properties of each function V_ϵ for fixed $\epsilon > 0$, and then we consider properties of the family $\{V_\epsilon\}_{\epsilon>0}$.

Many of the results in Section 9.1 correspond to similar results in Chapter 4 and Chapter 6 for the singular problem, for example we will see that V_ϵ is the unique viscosity solution of the penalized HJB equation in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$. The most important results in Section 9.2 are that V_ϵ is decreasing in ϵ, and that $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous on compact subsets of $[0,T) \times \overline{\mathscr{D}}$. These results will become useful when proving that V_ϵ converges uniformly to V on compact subsets of $[0,T) \times \overline{\mathscr{D}}$ in Chapter 10.

9.1 Properties of V_ϵ for fixed ϵ

In this section we consider properties of V_ϵ for fixed $\epsilon > 0$. Many proofs are similar to proofs found in Chapter 4 and Chapter 6, and will therefore not be included.

First we see that V_ϵ is well-defined, and satisfies the same growth and monotonicity properties as V. We will also see that V_ϵ satisfies Lemma 4.6, i.e., V_ϵ is bounded by a function $K(1 + x + y)^\gamma$. However, we will not state this result before in Section 9.2, as we will see that the constant K can be chosen independently of ϵ.

Lemma 9.1. *The value function V_ϵ is well-defined, non-decreasing in x and y, and is concave in x and y.*

Proof. See the proof of Lemmas 4.1, 4.4 and 4.5. □

We can also prove that V_ϵ satisfies the same regularity properties as V. We will generalize this result in the next section to a theorem concerning uniform *equi*continuity of the functions $\{V_\epsilon\}_{\epsilon>0}$.

Lemma 9.2 (Uniform continuity on bounded subsets). *The value function V_ϵ is uniformly continuous on bounded subsets of $\overline{\mathscr{D}_T}$, i.e., for each $\epsilon > 0$ and any bounded subset $\mathscr{O}_T$ of $\overline{\mathscr{D}_T}$, there exists a function ω_ϵ that is continuous at $(0,0,0)$, such that $\omega_\epsilon(0,0,0) = 0$ and*

$$V_\epsilon(t,x,y) - V_\epsilon(t',x',y') \le \omega_\epsilon(|t-t'|, |x-x'|, |y-y'|)$$

for all $(t,x,y), (t',x',y') \in \mathscr{O}_T$.

Proof. See the proof of Theorem 4.13. □

Lemma 9.3 (Uniform continuity in x and y). *The value function V_ϵ is uniformly continuous in x and y, i.e., there exists a function $\omega_\epsilon : \overline{\mathscr{D}_T} \to [0,\infty)$ that is continuous at $(0,0)$, such that $\omega_\epsilon(0,0) = 0$ and*

$$V_\epsilon(t,x,y) - V_\epsilon(t,x',y') \leq \omega_\epsilon(|x-x'|, |y-y'|)$$

for all $(x,y),(x',y') \in \overline{\mathscr{D}}$ and all $t \in [0,T]$.

Proof. See the proof of Lemma 4.9 and Theorem 4.10. □

Now we will develop viscosity solution theory for the penalized problem. As for the singular problem, we will see that the value function V_ϵ is the unique viscosity solution of the corresponding HJB equation in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$. The definition of viscosity solutions of (8.3) and (8.4) is similar as in the singular case.

Definition 9.4 (Viscosity solutions).

(1) Let $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$. Any function $v_\epsilon \in C(\overline{\mathscr{D}_T})$ is a viscosity subsolution (supersolution) of (8.3) *in $\mathscr{O}_T$ if and only if we have, for every $(t,X) \in \mathscr{O}_T$ and $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ such that (t,X) is a global maximum (minimum) relative to $\mathscr{O}_T$ of $v_\epsilon - \phi$,*

$$F_\epsilon\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big) \geq 0 \, (\leq 0). \tag{9.1}$$

(2) Any $v_\epsilon \in C(\overline{\mathscr{D}_T})$ is a constrained viscosity solution of (8.3) *if and only if v_ϵ is a viscosity supersolution of* (8.3) *in $\mathscr{D}_T$ and v is a viscosity subsolution of* (8.3) *in $[0,T) \times \overline{\mathscr{D}}$.*

(3) Any $v_\epsilon \in C(\overline{\mathscr{D}_T})$ is a constrained viscosity solution of the terminal value problem (8.3) *and* (8.4) *if and only if v_ϵ is a viscosity solution of* (8.3) *and v_ϵ satisfies* (8.4).

First we will prove that V_ϵ is a viscosity solution of (8.3) and (8.4). A proof is included, as it requires a slightly different approach than in the singular case.

Theorem 9.5. *The value function V_ϵ is a viscosity solution of the terminal value problem* (8.3) *and* (8.4) *for each $\epsilon > 0$.*

Proof. First we will show that V_ϵ is a viscosity supersolution on $\mathscr{D}_T$. Suppose $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$, and that $(t,X) \in \mathscr{D}_T$ is a global minimizer of $V_\epsilon - \phi$. By the exact same argument as in the proof of Theorem 6.14, we see that $G(D_X\phi) \leq 0$ and

$$F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big) \leq 0,$$

which implies that

$$\begin{aligned} F_\epsilon\big(t, X, &D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big) \\ &= F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big) + \frac{1}{\epsilon}\max\big\{G(D_X\phi);\, 0\big\} \\ &\leq 0. \end{aligned}$$

We see that V_ϵ is a viscosity supersolution in $\mathscr{D}_T$.

Now we will prove that V_ϵ is a viscosity subsolution in $[0,T) \times \overline{\mathscr{D}}$. Suppose $(t,X) \in [0,T) \times \overline{\mathscr{D}}$ is a maximum of $V_\epsilon - \phi$. By Lemmas 8.1 and 8.2, there exist a consumption strategy $(\pi^*, C^*) \in \mathscr{B}^\epsilon_{t,X}$ such that

$$V_\epsilon(t,X) = \mathrm{E}\left[\int_t^\tau e^{-\delta s} U\left(Y_s^{\pi^*,C^*}\right) ds + V_\epsilon\left(\tau, X_\tau^{\pi^*,C^*}, Y_\tau^{\pi^*,C^*}\right)\right]$$

for all $\tau \in [t,T]$. Since

$$(V_\epsilon - \phi)(t,X) \geq \mathrm{E}\left[(V_\epsilon - \phi)\left(\tau, X_\tau^{\pi^*,C^*}, Y_\tau^{\pi^*,C^*}\right)\right],$$

this implies that

$$\phi(t,X) \leq \mathrm{E}\left[\int_t^\tau e^{-\delta s} U\left(Y_s^{\pi^*,C^*}\right) ds + \phi\left(\tau, X_\tau^{\pi^*,C^*}, Y_\tau^{\pi^*,C^*}\right)\right].$$

Itô's formula now gives

$$\begin{aligned}
0 \leq \mathrm{E}\Bigg[\int_t^\tau & \phi_s\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) + e^{-\delta s} U\left(Y_s^{\pi^*,C^*}\right) - \beta Y_s^{\pi^*,C^*} \phi_x\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) \\
& + (\hat{r} + (\hat{\mu} - \hat{r})\pi_s) X_s^{\pi^*,C^*} \phi_x\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) \\
& + \frac{1}{2}(\sigma\pi X_s^{\pi^*,C^*})^2 \phi_{xx}\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) + \mathscr{I}^\pi\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}, \phi\right) \\
& + c_s^*\Big(\beta\phi_y\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) - \phi_x\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right)\Big) ds\Bigg] \\
\leq (\tau - t) \sup_{s\in[t,\tau]} \mathrm{E}\Bigg[& \phi_s\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) + e^{-\delta t} U\left(Y_s^{\pi^*,C^*}\right) \\
& - \beta Y_s^{\pi^*,C^*} \phi_x\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) + (\hat{r} + (\hat{\mu} - \hat{r})\pi_s) X_s^{\pi^*,C^*} \phi_x\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) \\
& + \frac{1}{2}(\sigma\pi X_s^{\pi^*,C^*})^2 \phi_{xx}(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}) + \mathscr{I}^\pi\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}, \phi\right) \\
& + c_s^*\Big(\beta\phi_y\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right) - \phi_x\left(s, X_s^{\pi^*,C^*}, Y_s^{\pi^*,C^*}\right)\Big)\Bigg].
\end{aligned}$$

Dividing by $\tau - t$, letting $\tau \to t^+$ and using that $X_s^{\pi^*,C^*}$ and $Y_s^{\pi^*,C^*}$ are right-continuous and that ϕ is smooth, we get

$$\begin{aligned}
0 \leq\ & \phi_t(t,X) + e^{-\delta t} U(y) - \beta y \phi_x(t,X) + (\hat{r} + (\hat{\mu} - \hat{r})\pi_s) x \phi_x(t,X) + \frac{1}{2}(\sigma\pi x)^2 \phi_{xx}(t,X) \\
& + \mathscr{I}^\pi(t,X,\phi) + c_t^*\big(\beta\phi_y(t,X) - \phi_x(t,X)\big) \\
\leq\ & \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big) + \frac{1}{\epsilon}\max\{G(D_X\phi); 0\}.
\end{aligned}$$

It follows that V_ϵ is a subsolution in $[0,T) \times \overline{\mathscr{D}}$. We see easily that V_ϵ satisfies (8.4), and that V_ϵ is a viscosity solution of (8.3) and (8.4). □

Our next goal is to prove uniqueness of viscosity solutions. The proof is based on a comparison principle. As in the singular case, we need a formulation of viscosity solutions using the integral operators $\mathscr{I}^\pi_\kappa$ and $\mathscr{I}^{\pi,\kappa}$, and we need a lemma similar to Lemma 6.12. These two results will be stated in Lemmas 9.6 and 9.7 below.

Lemma 9.6. *Let $v_\epsilon \in C_1(\overline{\mathscr{D}_T})$. Then v_ϵ is a subsolution (supersolution) of* (8.3) *if and only if we have, for every $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T})$ and $\kappa \in (0,1)$,*

$$\phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^{\pi,\kappa}(t, X, v_\epsilon, D_X\phi), \mathscr{I}_\kappa^\pi(t, X, \phi)\big) + \frac{1}{\epsilon}\max\big\{G(D_X\phi);\, 0\big\} \geq 0\,(\leq 0)$$

whenever $(t, X) \in [0,T) \times \overline{\mathscr{D}}$ ($(t, X) \in \mathscr{D}_T$) is a global maximum (minimum) relative to $[0,T) \times \overline{\mathscr{D}}$ ($\mathscr{D}_T$) of $v_\epsilon - \phi$.

Proof. See the proof of Lemma 6.4. □

Lemma 9.7. *Let $\underline{v_\epsilon} \in C_2(\overline{\mathscr{D}_T}) \cap C(\overline{\mathscr{D}_T})$ be a subsolution of* (8.3) *on $[0,T) \times \overline{\mathscr{D}}$, and let $\overline{v_\epsilon} \in C_2(\overline{\mathscr{D}_T}) \cap C(\overline{\mathscr{D}_T})$ be a supersolution of* (8.3) *on $\mathscr{D}_T$. Let $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T} \times \overline{\mathscr{D}_T})$ and $\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big) \in \big([0,T) \times \overline{\mathscr{D}}\big) \times \mathscr{D}_T$ be such that $\Phi : \overline{\mathscr{D}_T} \times \overline{\mathscr{D}_T} \to \mathbb{R}$ has a global maximum at $\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)$, where*

$$\Phi\big((t_1, X_1), (t_2, X_2)\big) := \underline{v_\epsilon}(t_1, X_1) - \overline{v_\epsilon}(t_2, X_2) - \phi\big((t_1, X_1), (t_2, X_2)\big).$$

Furthermore, assume that in a neighbourhood of $\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)$, there are continuous functions $g_0 : \mathbb{R}^6 \to \mathbb{R}$, $g_1, g_2 : \mathbb{R}^3 \to \mathbb{S}^3$ with $g_0\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big) > 0$ satisfying

$$D^2\phi \leq g_0\big((t_1, X_1), (t_2, X_2)\big) \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \begin{pmatrix} g_1(t_1, X_1) & 0 \\ 0 & g_2(t_2, X_2) \end{pmatrix}.$$

Then, for any $\varsigma \in (0,1)$ and $\kappa > 0$, there exist two matrices $A_1, A_2 \in \mathbb{S}^3$ satisfying

$$\begin{aligned} -\frac{2g_0\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)}{1-\varsigma} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} &\leq \begin{pmatrix} A_1 & 0 \\ 0 & -A_2 \end{pmatrix} - \begin{pmatrix} g_1(t_1^*, X_1^*) & 0 \\ 0 & g_2(t_2^*, X_2^*) \end{pmatrix} \\ &\leq \frac{g_0\big((t_1^*, X_1^*), (t_2^*, X_2^*)\big)}{\varsigma} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}, \end{aligned} \tag{9.2}$$

such that

$$\begin{aligned} 0 \leq \phi_t + F_\epsilon\big(t_1^*, X_1^*, D_X\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_1^*, X_1^*, \underline{v_\epsilon}, D_X\phi), \mathscr{I}_\kappa^\pi(t_1^*, X_1^*, \phi)\big) \\ + \frac{1}{\epsilon}\max\big\{G(D_X\phi);\, 0\big\} \end{aligned} \tag{9.3}$$

and

$$\begin{aligned} 0 \geq -\phi_t + F\big(t_2^*, X_2^*, -D_X\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_2^*, X_2^*, \overline{v_\epsilon}, -D_X\phi), \mathscr{I}_\kappa^\pi(t_2^*, X_2^*, -\phi)\big) \\ + \frac{1}{\epsilon}\max\big\{G(-D_X\phi);\, 0\big\}, \end{aligned} \tag{9.4}$$

where $\widehat{A}_1$ and $\widehat{A}_2$ denote the part of A_1 and A_2, respectively, associated with the variable X.

We do not need to work with *strict* supersolutions, as we did in the singular case. In the singular case we needed a strict supersolution in order to prove that

$$-\phi_t + F\big(t, X, -D_X\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t, X, \overline{v_\epsilon}, -D_X\phi), \mathscr{I}_\kappa^\pi(t, X, -\phi)\big) \leq 0,$$

where (t, X) is a global minimum of $\overline{v_\epsilon} - \phi$ for a supersolution $\overline{v_\epsilon}$ and some appropriately defined function ϕ. In the proof below this is not needed, as it follows immediately from the supersolution property of $\overline{v}$ that

$$0 \geq -\phi_t + F\big(t, X, -D_X\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t, X, \overline{v_\epsilon}, -D_X\phi), \mathscr{I}^{\pi}_{\kappa}(t, X, -\phi)\big) \\ + \frac{1}{\epsilon}\max\big\{G(-D_X\phi);\, 0\big\}.$$

We are now ready to prove a comparison principle. The following theorem corresponds to Theorem 6.15 in the singular case.

Theorem 9.8 (Comparison principle). *Assume $\underline{v_\epsilon} \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$ is a subsolution of (8.3) in $[0,T) \times \overline{\mathscr{D}}$, and that $\overline{v_\epsilon} \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$ is a supersolution of (8.3) in $\mathscr{D}_T$. Assume further that $\underline{v_\epsilon} \leq \overline{v_\epsilon}$ for $t = T$. Then $\underline{v_\epsilon} \leq \overline{v_\epsilon}$ everywhere on $\overline{\mathscr{D}_T}$.*

Proof. The proof follows the proof of Theorem 6.15 closely, and many details in this proof will therefore be skipped with reference to the proof of Theorem 6.15. Since $\underline{v}_\epsilon, \overline{v}_\epsilon \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$, there is a $\gamma' > 0$ such that $\underline{v}_\epsilon, \overline{v}_\epsilon \in C_{\gamma'}(\overline{\mathscr{D}_T})$ and $\delta > k(\gamma')$.

We know that

$$\underline{v_\epsilon}(t, X) - \overline{v_\epsilon}(t, X) \to -\infty$$

as $x, y \to \infty$, and that

$$\underline{v_\epsilon}(t, X) - \overline{v_\epsilon}(t, X) < 0$$

for $t = T$. When finding the maximum of $\underline{v_\epsilon}(t, X) - \overline{v_\epsilon}(t, X)$, it is therefore sufficient to consider the domain

$$\mathscr{O}_T := [0, T) \times \mathscr{O},$$

where

$$\mathscr{O} = \{(x, y) \in \mathscr{D} : 0 < x, y < R\}$$

for some appropriatly defined constant $R \in \mathbb{R}$. Assume $\underline{v_\epsilon} \leq \overline{v_\epsilon}$ is violated somewhere in $\mathscr{O}_T$, and define

$$M := \max_{\overline{\mathscr{O}_T}}(\underline{v_\epsilon} - \overline{v_\epsilon})(t, X) + \xi(t - T),$$

where $\xi > 0$ is chosen so small that $M > 0$. Assume $(t^*, X^*) \in \overline{\mathscr{O}_T}$ satisfies

$$M = (\underline{v_\epsilon} - \overline{v_\epsilon})(t^*, X^*) + \xi(t^* - T).$$

By choosing R and large enough, we can assume that we have either

1. $(t^*, X^*) \in \varGamma$, or
2. $x^*, y^* \in (0, R)$,

where $\varGamma = \{(t, X) \in \mathscr{O}_T : x = 0, y \in [0, R) \text{ or } x \in [0, R), y = 0\}$. We cannot have $t^* = T$, since $\underline{v_\epsilon} \leq \overline{v_\epsilon}$ for $t = T$.

Case 1: This is the case $(t^*, X^*) \in \varGamma$. We define $\alpha > 1$, $\epsilon' \in (0, 1)$, $h_0, \kappa_0 > 0$, $\eta : \overline{\mathscr{O}_T} \to \mathbb{R}^3$ and $\phi, \varPhi : \overline{\mathscr{O}_T} \times \overline{\mathscr{O}_T} \to \mathbb{R}$ exactly as in the proof of Theorem 6.15, i.e., we have

$$\mathscr{N}\big((t, X) + h\eta(t, X), h\kappa_0\big) \subset \mathscr{O}_T \quad \forall (t, X) \in \overline{\mathscr{O}_T} \text{ and } \forall h \in (0, h_0],$$

$$\phi\big((t_1, X_1), (t_2, X_2)\big) = \big|\alpha\big((t_1, X_1) - (t_2, X_2)\big) + \epsilon'\eta(t^*, X^*)\big|^2 \\ + \epsilon'\big|(t_1, X_1) - (t^*, X^*)\big|^2 - \xi(t_2 - T)$$

and

$$\varPhi\big((t_1, X_1), (t_2, X_2)\big) = \underline{v_\epsilon}(t_1, X_1) - \overline{v_\epsilon}(t_2, X_2) - \phi\big((t_1, X_1), (t_2, X_2)\big).$$

The purpose of the term $-\xi(t_2 - T)$, is to be make $\phi_{t_1} + \phi_{t_2}$ converge to something strictly negative when $\alpha \to \infty$ and $\epsilon' \to 0$. Define

$$M_\alpha := \max_{\overline{\mathscr{O}_T} \times \overline{\mathscr{O}_T}} \Phi\big((t_1, X_1), (t_2, X_2)\big) > 0,$$

and assume

$$M_\alpha = \Phi\big((t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha})\big)$$

for some $(t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha}) \in \overline{\mathscr{O}_T}$. As in the proof of Theorem 6.15, we see that

$$(t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha}) \to (t^*, X^*),$$
$$\alpha\big|(t_{1,\alpha}, X_{1,\alpha}) - (t_{2,\alpha}, X_{2,\alpha})\big| \to 0$$

and

$$M_\alpha \to M$$

as, in that order, $\alpha \to \infty$ and $\epsilon' \to 0$. Now we apply Lemma 9.7. Just as in the proof of Theorem 6.15, we see that there are matrices $A_1, A_2 \in \mathbb{S}^3$ such that

$$\begin{pmatrix} A_1 & 0 \\ 0 & -A_2 \end{pmatrix} \leq \frac{\alpha}{\varsigma} \begin{pmatrix} I & I \\ -I & I \end{pmatrix} + \epsilon' \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{aligned} 0 \leq \phi_{t_1} &+ F\big(t_{1,\alpha}, X_{1,\alpha}, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \underline{v}, D_{X_1}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi)\big) \\ &+ \frac{1}{\epsilon} \max\{G(D_{X_1}\phi); 0\} \end{aligned} \tag{9.5}$$

and

$$\begin{aligned} 0 \geq & -\phi_{t_2} + F\big(t_{2,\alpha}, X_{2,\alpha}, -D_{X_2}\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v}, -D_{X_2}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big) \\ &+ \frac{1}{\epsilon} \max\{G(-D_{X_2}\phi); 0\}, \end{aligned} \tag{9.6}$$

where ϕ_t and $D_X\phi$ are evaluated at $\big((t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha})\big)$ and $\widehat{A}_1, \widehat{A}_2$ are the parts of A_1, A_2 associated with X. Just as in the proof of Theorem 6.12, we have

$$\lim_{\epsilon \to 0} \lim_{\alpha \to \infty} \big(x_{1,\alpha}^2 a_{1,xx} - x_{2,\alpha}^2 a_{2,xx}\big) \leq 0.$$

Using (9.5) and (9.6), we get

$$\begin{aligned} 0 \leq \phi_{t_1} &+ F\Big(t_{1,\alpha}, X_{1,\alpha}, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \underline{v}, D_{X_1}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi)\Big) \\ &+ \frac{1}{\epsilon} \max\{G(D_{X_1}\phi); 0\} + \phi_{t_2} - F\Big(t_{2,\alpha}, X_{2,\alpha}, -D_{X_2}\phi, \widehat{A}_2, \\ &\mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v}, -D_{X_2}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\Big) - \frac{1}{\epsilon} \max\{G(-D_{X_2}\phi); 0\} \\ = &\big(e^{-\delta t_{1,\alpha}} U(y_{1,\alpha}) - e^{-\delta t_{2,\alpha}} U(y_{2,\alpha})\big) + \big(\phi_{t_1} + \phi_{t_2}\big) - \beta\big(y_{1,\alpha}\phi_{y_1} + y_{2,\alpha}\phi_{y_2}\big) \\ &+ \max_{\pi \in [0,1]} \Big[(\hat{r} + (\hat{\mu} - \hat{r})\pi)(x_{1,\alpha}\phi_{x_1} + x_{2,\alpha}\phi_{x_2}) + \frac{1}{2}(\sigma\pi)^2(x_{1,\alpha}^2 a_{1,xx} - x_{2,\alpha}^2 a_{2,xx}) \\ &+ \big(\mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v_\epsilon}, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v_\epsilon}, -D_{X_2}\phi)\big) \\ &+ \big(\mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big)\Big] \\ &+ \frac{1}{\epsilon}\Big(\max\{G(D_{X_1}\phi); 0\} - \max\{G(-D_{X_2}\phi); 0\}\Big). \end{aligned} \tag{9.7}$$

We have

$$\begin{aligned}
&\frac{1}{\epsilon}\Big(\max\{G(D_{X_1}\phi);\,0\}-\max\{G(-D_{X_2}\phi);\,0\}\Big)\\
&\qquad\le\frac{1}{\epsilon}\max\big\{G(D_{X_1}\phi)-G(-D_{X_2}\phi);\,0\big\}\\
&\qquad=\frac{\epsilon'}{\epsilon}\max\big\{2\beta(y_{1,\alpha}-y^*)-2(x_{1,\alpha}-x^*);\,0\big\}\\
&\qquad\to 0
\end{aligned}$$

as $\alpha\to\infty$. The other terms on the right-hand side of (9.7) also converge to something ≤ 0 as, in that order, $\alpha\to\infty$, $\epsilon'\to 0$ and $\kappa\to 0$, see the proof of Theorem 6.15. The term $\phi_{t_1}+\phi_{t_2}$ converges to $-\xi$, so the right-hand side of (9.7) converges to something strictly negative. The have obtained a contradiction, since the left-hand side of (9.7) is 0.

Case 2: Exactly as in Case 2 in the proof of Theorem 6.15, we define $\phi,\varPhi:\overline{\mathscr{O}_T}\times\overline{\mathscr{O}_T}\to\mathbb{R}$ by

$$\phi\big((t_1,X_1),(t_2,X_2)\big)=\frac{\alpha}{2}\big|(t_1,X_1)-(t_2,X_2)\big|^2-\xi(t_2-T)$$

and

$$\varPhi\big((t_1,X_1),(t_2,X_2)\big)=\underline{v_\epsilon}(t_1,X_1)-\overline{v_\epsilon}(t_2,X_2)-\phi\big((t_1,X_1),(t_2,X_2)\big).$$

Now define

$$M_\alpha=\max_{\overline{\mathscr{O}_T}\times\overline{\mathscr{O}_T}}\varPhi\big((t_1,X_1),(t_2,X_2)\big)>0,$$

and suppose

$$M_\alpha=\varPhi\big((t_{1,\alpha},X_\alpha),(t_{2,\alpha},X_{2,\alpha})\big),$$

for some $(t_{1,\alpha},X_\alpha),(t_{2,\alpha},X_{2,\alpha})\in\overline{\mathscr{O}_T}$. By the result in Case 1, we see that any limit point of $(t_{1,\alpha},X_\alpha)$ and $(t_{2,\alpha},X_{2,\alpha})$ must belong to $\mathscr{O}_T$. We know that $M_\alpha>0$, because $M_\alpha\ge M>0$. As in the proof of Theorem 6.15, we see that

$$\alpha|(t_{1,\alpha},X_{1,\alpha})-(t_{2,\alpha},X_{2,\alpha})|\to 0$$

and that $M_\alpha\to M$ as $\alpha\to\infty$. By Lemma 9.7 with assumptions (1)-(5), we see that, for any $\varsigma\in(0,1)$, there are matrices $A_1,A_2\in\mathbb{S}^3$ such that

$$\begin{pmatrix}A_1&0\\0&A_2\end{pmatrix}\le\frac{\alpha}{\varsigma}\begin{pmatrix}I&-I\\-I&I\end{pmatrix},$$

$$\begin{aligned}
0\le\,&\phi_{t_1}+F\Big(t_{1,\alpha},X_{1,\alpha},D_{X_1}\phi,\widehat{A}_1,\mathscr{I}^{\pi,\kappa}(t_{1,\alpha},X_{1,\alpha},\underline{v},D_{X_1}\phi),\mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha},X_{1,\alpha},\phi)\Big)\\
&+\frac{1}{\epsilon}\max\{G(D_{X_1}\phi);\,0\}
\end{aligned}\tag{9.8}$$

and

$$\begin{aligned}
0\ge\,&-\phi_{t_2}+F\Big(t_{2,\alpha},X_{2,\alpha},-D_{X_2}\phi,\widehat{A}_2,\mathscr{I}^{\pi,\kappa}(t_{2,\alpha},X_{2,\alpha},\underline{v},-D_{X_2}\phi),\mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha},X_{2,\alpha},-\phi)\Big)\\
&+\frac{1}{\epsilon}\max\{G(-D_{X_2}\phi);\,0\},
\end{aligned}\tag{9.9}$$

where ϕ_{t_1}, $D_{X_1}\phi$, ϕ_{t_2} and $D_{X_2}\phi$ are evaluated at $\big((t_{1,\alpha}, X_{1,\alpha}), (t_{2,\alpha}, X_{2,\alpha})\big)$. As in Case 1, we get

$$\lim_{\alpha\to\infty}(x_{1,\alpha}^2 a_{1,xx} - x_{2,\alpha}^2 a_{2,xx}) \le 0.$$

Equations (9.8) and (9.9) imply that

$$\begin{aligned}
0 \le\; &\phi_{t_1} + F\big(t_{1,\alpha}, X_{1,\alpha}, D_{X_1}\phi, \widehat{A}_1, \mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v_\epsilon}, D_{X_1}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi)\big)\\
&+ \phi_{t_2} - F\big(t_{2,\alpha}, X_{2,\alpha}, -D_{X_2}\phi, \widehat{A}_2, \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \overline{v_\epsilon}, -D_{X_2}\phi), \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\big)\\
\le\; &\big(e^{-\delta t_{1,\alpha}}U(y_{1,\alpha}) - e^{-\delta t_{2,\alpha}}U(y_{2,\alpha})\big) + \big(\phi_{t_1} + \phi_{t_2}\big) - \beta\big(y_{1,\alpha}\phi_{y_1} + y_{2,\alpha}\phi_{y_2}\big)\\
&+ \max_{\pi\in[0,1]}\Big[(\hat{r} + (\hat{\mu} - \hat{r})\pi)(x_{1,\alpha}^2\phi_{x_1} - x_{2,\alpha}^2\phi_{t_2}) + \frac{1}{2}(\sigma\pi)^2(x_{1,\alpha}^2 a_{1,xx} - x_{2,\alpha}^2 a_{2,xx})\\
&+ \Big(\mathscr{I}^{\pi,\kappa}(t_{1,\alpha}, X_{1,\alpha}, \overline{v_\epsilon}, D_{X_1}\phi) - \mathscr{I}^{\pi,\kappa}(t_{2,\alpha}, X_{2,\alpha}, \underline{v_\epsilon}, -D_{X_2}\phi)\Big)\\
&+ \Big(\mathscr{I}^{\pi}_{\kappa}(t_{1,\alpha}, X_{1,\alpha}, \phi) - \mathscr{I}^{\pi}_{\kappa}(t_{2,\alpha}, X_{2,\alpha}, -\phi)\Big)\Big]\\
&+ \frac{1}{\epsilon}\Big(\max\{G(D_{X_1}\phi); 0\} - \max\{G(-D_{X_2}\phi); 0\}\Big).
\end{aligned} \tag{9.10}$$

By letting, in that order, $\alpha \to \infty$ and $\kappa \to \infty$ in (9.10), we get a contradiction, as the right-hand side converges to something < 0. This concludes the proof of the theorem. □

The comparison principle implies uniqueness of viscosity solutions, just as in the singular case.

Theorem 9.9 (Uniqueness of viscosity solutions in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$). *Viscosity solutions of the terminal value problem* (8.3) *and* (8.4) *are unique in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$ for each $\epsilon > 0$.*

Proof. See the proof of Theorem 6.16. □

9.2 Properties of the family $\{V_\epsilon\}_{\epsilon>0}$

In this section we consider properties of the family $\{V_\epsilon\}_{\epsilon>0}$. The purpose of the section is to establish properties that will help us when we prove convergence of V_ϵ to V.

The two main results are Theorem 9.11 and Theorem 9.18. Theorem 9.11 says that V_ϵ increases as ϵ decreases, and this property will be used many times both in this section and in Chapter 10. Theorem 9.18 says that $\{V_\epsilon\}_{\epsilon>0}$ is uniformly equicontinuous on compact subsets of $[0,T) \times \overline{\mathscr{D}}$, and this property will be used in Sections 10.3 and 10.4 when we prove convergence of V_ϵ by the Arzelà-Ascoli theorem.

Lemma 9.10. *We have $\mathscr{B}^{\epsilon_1}_{t,x,y} \subset \mathscr{B}^{\epsilon_2}_{t,x,y} \subset \mathscr{A}_{t,x,y}$ for all $(t,x,y) \in \overline{\mathscr{D}_T}$ and $\epsilon_1, \epsilon_2 > 0$ with $\epsilon_1 > \epsilon_2$.*

Proof. This follows directly from the definition of $\mathscr{B}^{\epsilon}_{t,x,y}$, see equation (8.2). □

The following theorem says that V_ϵ increases when ϵ decreases, and, though its proof is simple, it will turn out to be extremely useful in both this and the next sections.

Theorem 9.11 (Monotonicity in ϵ). *Assume $(t,x,y)\in\overline{\mathscr{D}_T}$ and $\epsilon_1,\epsilon_2>0$, where $\epsilon_2<\epsilon_1$. Then we have $V_{\epsilon_1}(t,x,y)\le V_{\epsilon_2}(t,x,y)\le V(t,x,y)$.*

Proof. By Lemma 9.10, we have

$$\begin{aligned}V_\epsilon(t,x,y) &= \sup_{(\pi,C)\in\mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds + W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\right]\\ &\le \sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds + W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\right]\\ &= V(t,x,y)\end{aligned}$$

for all $\epsilon>0$. This shows that $V_\epsilon(t,x,y)\le V(t,x,y)$ for $\epsilon=\epsilon_1,\epsilon_2$. To show that $V_{\epsilon_1}(t,x,y)\le V_{\epsilon_2}(t,x,y)$, we note that

$$\begin{aligned}V_{\epsilon_1}(t,x,y) &= \sup_{(\pi,C)\in\mathscr{B}^{\epsilon_1}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds + W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\right]\\ &= \sup_{(\pi,C)\in\mathscr{B}^{\epsilon_2}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds + W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\right]\\ &= V_{\epsilon_2}(t,x,y),\end{aligned}$$

again by Lemma 9.10. □

The following lemma says that $\{V_\epsilon\}_{\epsilon>0}$ is uniformly bounded on compact subsets of $\overline{\mathscr{D}_T}$, and corresponds to Lemma 4.6 in the singular case.

Lemma 9.12. *There is a constant $K>0$ such that $0\le V_\epsilon(t,x,y)\le K(1+x+y)^\gamma$ for all $(t,x,y)\in\overline{\mathscr{D}_T}$ and $\epsilon>0$.*

Proof. This follows directly from Theorem 4.6, Theorem 9.11, and the fact that U and W take non-negative values. □

The following lemma corresponds to Lemma 4.8 in the singular case. It will be applied when we prove left-equicontinuity of $\{V_\epsilon\}_{\epsilon>0}$, and when we prove that $\lim_{\epsilon\to0}V_\epsilon(t,x,y)$ converges to $V(T,x,y)$ when $t\to T$.

Lemma 9.13. *For any $(t,x,y)\in\overline{\mathscr{D}_T}$ with $t\ne T$,*

$$\lim_{\epsilon\to0}\left(V_\epsilon(t,x,y)-\max_{c\in[0,x]}V_\epsilon(t,x-c,y+\beta c)\right)=0. \tag{9.11}$$

The convergence is uniform on compact subsets of $[0,T)\times\overline{\mathscr{D}}$.

Proof. The proof will consist of two parts: First we show that (9.11) holds for all $(t,x,y)\in\overline{\mathscr{D}_T}$ with $t\ne T$ (part 1). Then we prove that the convergence is uniform on compact subsets of $[0,T)\times\overline{\mathscr{D}}$ (part 2).

Part 1: Fix $(t,x,y)\in[0,T)\times\overline{\mathscr{D}}$. We see immediately that the limit on the left-hand side of (9.11) must be ≤ 0, since

$$V_\epsilon(t,x,y)=V_\epsilon(t,x-0,y+\beta\cdot0)\le\max_{c\in[0,x]}V_\epsilon(t,x-c,y+\beta c)$$

for all $\epsilon > 0$.

We will now prove that the limit on the left-hand side of (9.11) is ≥ 0. Let $c \in [0, x]$ be the maximizing value of $V_\epsilon(t, x-c, y+\beta c)$, and fix any $\epsilon' > 0$. The processes having initial values $x - c$ and $y + \beta c$ are denoted by $X_s'^{\pi',C'}$ and $Y_s'^{\pi',C'}$, respectively, and for each $\epsilon > 0$, assume $(\pi'_\epsilon, C'_\epsilon) \in \mathscr{B}^\epsilon_{t,x-c,y+\beta c}$ is an optimal control. We want to prove that there is an $\epsilon^* > 0$ such that

$$V_\epsilon(t, x, y) - V_\epsilon(t, x-c, y+\beta c) \geq g(\epsilon') \tag{9.12}$$

for all $\epsilon \in (0, \epsilon^*)$, where g is some function converging to 0 as $\epsilon' \to 0$.

Denote the processes with initial value x and y by $X_s^{\pi_\epsilon,C_\epsilon}$ and $Y_s^{\pi_\epsilon,C_\epsilon}$, respectively. For each, $\epsilon > 0$ $(\pi_\epsilon, C_\epsilon) \in \mathscr{B}^\epsilon_{t,x,y}$ is defined as follows:

1. $(\pi_\epsilon)_s = (\pi_\epsilon)'_s$ for all $s \in [t, T]$, and
2. $(C_\epsilon)_s$ has gradient $1/\epsilon$ in the interval $[t, s']$, where $s' \geq t$ is defined to be the smallest time $s \in [t, T]$ that gives us $X_s^{\pi_\epsilon,C_\epsilon} = X_s'^{\pi'_\epsilon,C'_\epsilon}$. For $s > s'$, define $d(C_\epsilon)_s = d(C'_\epsilon)_s$. If $X_s^{\pi_\epsilon,C_\epsilon} > X_s'^{\pi'_\epsilon,C'_\epsilon}$ for all $s \in [t, T]$, we define $s' = \infty$.

Note that both π_ϵ and C_ϵ are adapted processes, and that s' is well-defined, because the paths of $X_s^{\pi_\epsilon,C_\epsilon}$ and $X_s'^{\pi'_\epsilon,C'_\epsilon}$ are right-continuous.

Now define $\epsilon^* > 0$ to be so small that

(1) $s' - t < \epsilon'$,
(2) $|(C_\epsilon)_{s'} - c - (C'_\epsilon)_{s'}| < \epsilon'$, and
(3) $t + s' < T$

of probability at least $1 - \epsilon'$ for all $\epsilon < \epsilon^*$. We need to prove that such a value of ϵ^* exists, and we will divide the proof into two parts: First we prove that an ϵ^* satisfying (1)-(3) almost certainly exists if we fix $\omega \in \Omega$, i.e., if we let the stochastic development of the Lévy process L_t be fixed (step 1). Then we will prove that there exists an $\epsilon^* > 0$ such that (1)-(3) are satisfied of probability at least $1 - \epsilon'$ (step 2).

Step 1: Fix $\omega \in \Omega$. We want to prove that, almost certainly, there is an $\epsilon^* > 0$ such that (1)-(3) are satisfied for all $\epsilon < \epsilon^*$. We see that $s' \to 0$ as $\epsilon \to 0$, so (1) and (3) can easily be satisfied. We know that $X_{s'}^{\pi_\epsilon,C_\epsilon} = X_{s'}'^{\pi'_\epsilon,C'_\epsilon}$, and therefore (3.3) implies

$$\begin{aligned}
|C_{s'} - c - C'_{s'}| &= \int_t^{s'} (\hat{r} + (\hat{\mu} - \hat{r})\pi_{s''}) \left(X_{s''}^{\pi_\epsilon,C_\epsilon} - X_{s''}'^{\pi'_\epsilon,C'_\epsilon}\right) ds'' \\
&\quad + \int_t^{s'} \sigma(\pi_\epsilon)_{s''} \left(X_{s''}^{\pi_\epsilon,C_\epsilon} - X_{s''}'^{\pi'_\epsilon,C'_\epsilon}\right) dB_{s''} \\
&\quad + \int_t^{s'} (\pi'_\epsilon)_{s''} \left(X_{s''}^{\pi_\epsilon,C_\epsilon} - X_{s''}'^{\pi'_\epsilon,C'_\epsilon}\right) \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)\, \tilde{N}(ds'', dz) \\
&\to 0
\end{aligned}$$

almost certainly as $s' \to 0$, so (2) can also be satisfied. It follows that we almost certainly manage to choose an $\epsilon^* > 0$ such that (1)-(3) are satisfied all $\epsilon < \epsilon^*$.

Step 2: Now we want to prove that there is an $\epsilon^* > 0$ such that (1)-(3) are satisfied with probability at least $1 - \epsilon'$ for all $\epsilon < \epsilon^*$. Define the set $\Omega_\epsilon \subset \Omega$ to be such that $\omega \in \Omega_\epsilon$ iff ω satisfies (1)-(3). We see easily that $\Omega_{\epsilon_1} \subset \Omega_{\epsilon_2}$ for $\epsilon_2 < \epsilon_1$. By the result in step 1, almost all $\omega \in \Omega$ belong to Ω_ϵ for sufficiently small $\epsilon > 0$. Therefore it is possible to choose an $\epsilon^* > 0$ such that $\mathbb{P}(\Omega_\epsilon) \geq 1-\epsilon'$ for all $\epsilon < \epsilon^*$. This completes the second part of the proof, and the existence of an appropriate $\epsilon^* > 0$ is proved.

Assume $\epsilon < \epsilon^*$, and denote $(\pi_\epsilon, C_\epsilon)$ and $(\pi'_\epsilon, C'_\epsilon)$ by (π, C) and (π', C'), respectively, from now on. We want to prove (9.12). We see that $X_s^{\pi,C} = X_s'^{\pi',C'}$ for $s \geq s'$, because $X_{s'}^{\pi,C} = X_{s'}'^{\pi',C'}$, and $dC_s = dC'_s$ and $\pi_s = \pi'_s$ for $s > s'$. For $s < s'$, we have $X_s^{\pi,C} \geq X_s'^{\pi',C'}$, because $X_t^{\pi,C} \geq X_t'^{\pi',C'}$.

For $s \geq s'$ we have

$$\begin{aligned}
&Y_s^{\pi,C} - Y_s'^{\pi',C'}\\
&\quad= \Big(ye^{-\beta(s-t)} + \beta e^{-\beta s}\int_t^{s'} e^{\beta s''}\,dC_{s''}\Big) - \Big((y+\beta c)e^{-\beta(s-t)} + \beta e^{-\beta s}\int_t^{s} e^{\beta s''}dC'_{s''}\Big)\\
&\quad= -\beta c e^{-\beta(s-t)} + \beta e^{-\beta s}\int_t^{s} e^{\beta s''}(dC_{s''} - dC'_{s''})\\
&\quad\geq -\beta c e^{-\beta(s-t)} + \beta e^{-\beta(s-t)}\int_t^{s'} (dC_{s''} - dC'_{s''})\\
&\quad= \beta e^{-\beta(s-t)}(-c + C_{s'} - C'_{s'}),
\end{aligned}$$

and for $s < s'$ we have

$$\begin{aligned}
&Y_s^{\pi,C} - Y_s'^{\pi',C'}\\
&\quad= \Big(ye^{-\beta(s-t)} + \beta e^{-\beta s}\int_t^{s} e^{\beta s''}dC_{s''}\Big) - \Big((y+\beta c)e^{-\beta(s-t)} + \beta e^{-\beta s}\int_t^{s} e^{\beta s''}dC'_{s''}\Big)\\
&\quad\geq -\beta c e^{-\beta(s-t)},
\end{aligned}$$

because $dC_{s''} \geq dC'_{s''}$ for $s'' \in [t,s']$. Using the estimates for $X_s^{\pi,C} - X_s'^{\pi',C'}$ and $Y_s^{\pi,C} - Y_s'^{\pi',C'}$ above, we see that

$$\begin{aligned}
&\int_t^T e^{-\delta s}\Big(U\left(Y_s^{\pi,C}\right) - U\left(Y_s'^{\pi',C'}\right)\Big)\,ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right) - W\left(X_T'^{\pi',C'}, Y_T'^{\pi',C'}\right)\\
&\quad\geq -(s'-t)e^{-\beta t}\omega_U\left(\beta c e^{-\beta(s-t)}\right) - (T-s')e^{-\beta s'}\omega_U\left(\beta e^{-\beta(s-t)}(-c + C_{s'} - C'_{s'})\right)\\
&\qquad - \omega_W\left(0, \beta e^{-\beta(s-t)}(-c + C_{s'} - C'_{s'})\right)
\end{aligned}$$

if $s' < T$, and

$$\begin{aligned}
&\int_t^T e^{-\delta s}\Big(U\left(Y_s^{\pi,C}\right) - U\left(Y_s'^{\pi',C'}\right)\Big)\,ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right) - W\left(X_T'^{\pi',C'}, Y_T'^{\pi',C'}\right)\\
&\qquad\qquad \geq -(T-t)\omega_U\left(\beta c e^{-\beta(s-t)}\right) - \omega_W\left(0, \beta c e^{-\beta(s-t)}\right)
\end{aligned}$$

for $s' = \infty$, where ω_U and ω_W are moduli of continuity for U and W, respectively.

If (1)-(3) are satisfied,

$$\begin{aligned}
&\int_t^T e^{-\delta s}\Big(U\left(Y_s^{\pi,C}\right) - U\left(Y_s'^{\pi',C'}\right)\Big)\,ds + W\left(X_T^{\pi,C}, Y_T^{\pi,C}\right) - W\left(X_T'^{\pi',C'}, Y_T'^{\pi',C'}\right)\\
&\qquad \geq -\epsilon' e^{-\beta t}\omega_U\big(\beta c e^{-\beta(s-t)}\big) - (T-s')e^{-\beta s'}\omega_U\big(\epsilon'\beta e^{-\beta(s-t)}\big)\\
&\qquad\quad - \omega_W\big(0, \epsilon'\beta e^{-\beta(s-t)}\big),
\end{aligned}$$

and if (1)-(3) are not satisfied,

$$\begin{aligned}\int_t^T & e^{-\delta s}\Big(U\left(Y_s^{\pi,C}\right)-U\left(Y_s^{\prime\pi',C'}\right)\Big)\,ds+W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)-W\left(X_T^{\prime\pi',C'},Y_T^{\prime\pi',C'}\right)\\ &\geq -(T-t)\beta c e^{-\beta(s-t)}-\omega_W\left(0,\beta c e^{-\beta(s-t)}\right),\end{aligned}$$

provided ϵ' is sufficiently small.

Since (1)-(3) are satisfied with probability at least $1-\epsilon'$, we get

$$\begin{aligned}V_\epsilon&(t,x,y)-V_\epsilon(t,x-c,y+\beta c)\\ &\geq \mathrm{E}\Big[\int_t^T e^{-\delta s}\Big(U\left(Y_s^{\pi,C}\right)-U\left(Y_s^{\prime\pi',C'}\right)\Big)\,ds+W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\\ &\quad -W\left(X_T^{\prime\pi',C'},Y_T^{\prime\pi',C'}\right)\Big]\\ &\geq -\epsilon' e^{-\beta t}\omega_U\left(\beta c e^{-\beta(s-t)}\right)-(T-s')e^{-\beta s'}\omega_U\left(\epsilon'\beta e^{-\beta(s-t)}\right)\\ &\quad -\omega_W\left(0,\epsilon'\beta c e^{-\beta(s-t)}\right)+\epsilon'\Big(-(T-t)\beta c e^{-\beta(s-t)}-\omega_W\left(0,\beta c e^{-\beta(s-t)}\right)\Big).\end{aligned} \tag{9.13}$$

We see that the right-hand side of this inequality can be written as $g(\epsilon')$, where $g(\epsilon')\to 0$ as $\epsilon'\to 0$, so (9.12) is proved. The estimate (9.11) follows.

Part 2: We will now prove that the convergence given by (9.11) is uniform on compact subsets of $[0,T)\times\overline{\mathscr{D}}$. Let $\mathscr{O}_T$ be a compact subset of $[0,T)\times\overline{\mathscr{D}}$. We need to prove that both ϵ^* and the function g given above, can be defined independently of $(t,x,y)\in\mathscr{O}_T$.

First we will prove that ϵ^* can be defined independently of (t,x,y), i.e., there is an $\epsilon^*>0$, such that (1)-(3) hold of probability at least $1-\epsilon'$ for all $\epsilon<\epsilon^*$, uniformly in (t,x,y). We divide the proof into two steps as above:

Step 1: Fix $\omega\in\Omega$. Since x is bounded, we see that (1) and (2) hold uniformly in (t,x,y) for sufficiently small ϵ^*. Since t is bounded uniformly away from T, we see that (3) holds uniformly in (t,x,y) for sufficiently small ϵ^*. It follows that (1)-(3) holds uniformly in (t,x,y) for fixed $\omega\in\Omega$ and sufficiently small ϵ^*.

Step 2: This step can be done just in Part 1. We conclude that ϵ^* can be defined independently of (t,x,y).

The right-hand side of (9.13) converges uniformly to 0 when $\epsilon'\to 0$, because $c<x$ is bounded. We conclude that also g can be chosen independently of (t,x,y), and the result of the lemma follows. □

Our next goal is to prove Theorem 9.18, i.e., $\{V_\epsilon\}_{\epsilon>0}$ is uniformly equicontinuous on compact subsets of $[0,T)\times\overline{\mathscr{D}}$.

The following theorem shows that $\{V_\epsilon\}_{\epsilon>0}$ is uniformly equicontinuous in x and y. The proof is based on the fact that V_ϵ is bounded by V (Theorem 9.11), and that the result holds for V (Theorem 4.10).

Theorem 9.14 (Uniform equicontinuity in x and y). *The set of functions $\{V_\epsilon\}_{\epsilon>0}$ is uniformly equicontinuous in x and y on $\overline{\mathscr{D}}$ i.e., there exists a function $\omega:\overline{\mathscr{D}}\to\mathbb{R}$, such that $\omega(0,0)=0$, ω is continuous at $(0,0)$, and*

$$|V_\epsilon(t,x',y')-V_\epsilon(t,x,y)|\leq\omega(|x'-x|,|y'-y|)$$

for all $(x,y),(x',y')\in\overline{\mathscr{D}}$, $t\in[0,T]$ and $\epsilon>0$.

Proof. Let $\epsilon > 0$. Proceeding exactly as in the proof of Theorems 4.9 and 4.10, we see that

$$|V_\epsilon(t,x,y) - V(t,x',y')_\epsilon| \leq \omega_{t,\epsilon}(|x-x'|,|y-y'|),$$

where

$$\omega_{t,\epsilon}(x,y) := \sup_{(\pi,C)\in\mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}\omega_U\left(Y_s^{\pi,C}\right)\,ds + \omega_W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\right],$$

and ω_U and ω_W are moduli of continuity for U and W, respectively. By Theorem 9.11, we know that

$$\omega_{t,\epsilon}(x,y) \leq \omega_t(x,y)$$

for all $(t,x,y) \in \overline{\mathscr{D}_T}$, where

$$\omega_t(x,y) := \sup_{(\pi,C)\in\mathscr{A}_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s}\omega_U\left(Y_s^{\pi,C}\right)\,ds + \omega_W\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\right].$$

The function $\omega : \overline{\mathscr{D}} \to \mathbb{R}$ defined in the proofs of Theorem 4.10 is continuous at $(0,0)$,

$$\omega_t(x,y) \leq \omega(x,y)$$

for all $t \in [0,T]$, and $\omega(0,0) = 0$. We see that ω satisfies all the desired properties. □

We will now show that $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous in t for $t < T$. We split this into the cases of right-equicontinuity and left-equicontinuity, and each proof is again divided into upper semi-continuity and lower semi-continuity. The proofs are based on the dynamic programming principle (Theorem 8.1), the fact that V_ϵ increases as ϵ decreases (Theorem 9.11), that each function V_ϵ is continuous in t (Lemma 9.2), and that $\{V_\epsilon\}_{\epsilon>0}$ is uniformly equicontinuous in x and y (Lemma 9.14).

First we will define what we mean by right semi-continuity and left semi-continuity. These two terms will be used in the proofs below. The definition we give here is consistent with definitions found in other articles and books, see for example [28].

Definition 9.15. *Let $f : A \to \mathbb{R}$, where $A \subset \mathbb{R}$. We say that the function f is upper semi-continuous from the right (left) if, given any $\epsilon > 0$ and $a^* \in A$, there is a $\delta > 0$, such that*

$$f(a) - f(a^*) < \epsilon,$$

for all $a \in [a^*, a^*+\delta)$ $\big(a \in (a^*-\delta, a^*]\big)$. We say that f is lower semi-continuous from the right (left) if, given any $\epsilon > 0$ and $a^* \in A$, there is a $\delta > 0$, such that

$$f(a) - f(a^*) > -\epsilon,$$

for all $a \in [a^*, a^*+\delta)$ $\big(a \in (a^*-\delta, a^*]\big)$.

Lemma 9.16. *$\{V_\epsilon\}_{\epsilon>0}$ is right-equicontinuous in t on $\overline{\mathscr{D}_T}$, i.e., for each $(t,x,y) \in \overline{\mathscr{D}_T}$, there is a function ω such that $|V_\epsilon(t,x,y) - V_\epsilon(t',x,y)| \leq \omega(|t-t'|)$ for all $t' \in (t,T]$, ω is continuous at 0, and $\omega(0) = 0$.*

Proof. First we prove that $\{V_\epsilon\}_{\epsilon>0}$ is upper semi-continuous from the right, uniformly in ϵ. Fix $(t,x,y) \in \overline{\mathscr{D}_T}$, $t \neq T$. For all $\epsilon > 0$ and $t' \in [t,T]$ we have

$$\begin{aligned}
V_\epsilon(t,x,y) = & \sup_{(\pi,C)\in\mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_t^{t'} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\, ds + V_\epsilon\left(t', X_{t'}^{\pi,C}, Y_{t'}^{\pi,C}\right)\right] \\
\geq & \,\mathrm{E}\left[V_\epsilon\left(t', X_{t'}^{0,0}, Y_{t'}^{0,0}\right)\right] \\
= & \,V_\epsilon\left(t', x e^{\hat{r}(t'-t)}, y e^{-\beta(t'-t)}\right) \\
\geq & \,V_\epsilon(t',x,y) - \hat{\omega}\left(x\left(e^{\hat{r}(t'-t)} - 1\right), y\left(e^{-\beta(t'-t)} - 1\right)\right),
\end{aligned}$$

where $\hat{\omega}$ is a modulus of continuity for V_ϵ in x and y for all $t \in [0,T]$ and $\epsilon > 0$. The existence of $\hat{\omega}$ follows from Theorem 9.14. The explicit expressions for $X_{t'}^{0,0}$ and $Y_{t'}^{0,0}$ in the last line follow from (3.2) and (3.3). Letting $t' \to t^+$, we see that $\{V_\epsilon\}_{\epsilon>0}$ is upper semi-continuous from the right, uniformly in ϵ.

Now we will to prove that $\{V_\epsilon\}_{\epsilon>0}$ is lower semi-continuous from the right, uniformly in ϵ. Suppose the contrary, i.e., there are sequences $\{\epsilon_n\}_{n\in\mathbb{N}}$, $\{t_n\}_{n\in\mathbb{N}}$ and an $\epsilon' > 0$, such that $t_n \to t^+$ and

$$V_{\epsilon_n}(t,x,y) - V_{\epsilon_n}(t_n,x,y) > \epsilon' \tag{9.14}$$

for all $n \in \mathbb{N}$.

First suppose 0 is a limit point of $\{\epsilon_n\}$. Then we can suppose without loss of generality that $\{\epsilon_n\}$ and $\{t_n\}$ are decreasing, and that $\epsilon_n \to 0$. By Theorem 9.11, $\{V_{\epsilon_n}(t,x,y)\}_{n\in\mathbb{N}}$ is an increasing sequence that is bounded above. Therefore it is convergent, and we define $a = \lim_{n\to\infty} V_{\epsilon_n}(t,x,y)$. There is an $N \in \mathbb{N}$ such that $a - V_{\epsilon_n}(t,x,y) < \epsilon'/2$ for all $n \geq N$. Since V_{ϵ_N} is continuous in t (Theorem 9.14) and $t_n \to t$, there is an $M \geq N$ such that $|V_{\epsilon_N}(t,x,y) - V_{\epsilon_N}(t_M,x,y)| < \epsilon'/2$. Theorem 9.11 says that $V_{\epsilon_M}(t,x,y) \leq a$ and $V_{\epsilon_M}(t_M,x,y) \geq V_{\epsilon_N}(t_M,x,y)$. Combining the inequalities we have obtained, we get

$$\begin{aligned}
V_{\epsilon_M}(t,x,y) - V_{\epsilon_M}(t_M,x,y) &\leq a - V_{\epsilon_N}(t_M,x,y) \\
&= \Big(a - V_{\epsilon_N}(t,x,y)\Big) + \Big(V_{\epsilon_N}(t,x,y) - V_{\epsilon_N}(t_M,x,y)\Big) \\
&< \epsilon'/2 + \epsilon'/2 \\
&= \epsilon',
\end{aligned}$$

which is a contradiction to (9.14).

Now suppose 0 is not a limit point of $\{\epsilon_n\}$. This case is necessary to consider, as we are not only proving equicontinuity of $\{V_{\epsilon_m}\}_{m\in\mathbb{N}}$ for some sequence $\{\epsilon_m\}_{m\in\mathbb{N}}$ converging to 0, but are proving equicontinuity of the whole family $\{V_\epsilon\}_{\epsilon>0}$. If 0 is not a limit point of $\{\epsilon_n\}$, there is an $\epsilon^* > 0$ such that $\epsilon_n > \epsilon^*$ for all $n \in \mathbb{N}$. By the dynamic programming principle (Theorem 8.1), we have

$$
\begin{aligned}
V_{\epsilon_n}&(t,x,y) - V_{\epsilon_n}(t_n,x,y)\\
&= \sup_{(\pi,C)\in\mathscr{B}^{\epsilon_n}_{t,x,y}} \mathrm{E}\left[\int_t^{t_n} e^{-\delta s} U\left(Y_s^{\pi,C}\right)\,ds + V\left(t_n, X_{t_n}^{\pi,C}, Y_{t_n}^{\pi,C}\right)\right]\\
&\quad - V_{\epsilon_n}(t_n,x,y)\\
&\le \sup_{(\pi,0)\in\mathscr{B}^{\epsilon_n}_{t,x,y}} \mathrm{E}\bigg[\int_t^{t_n} e^{-\delta s} U\left(y+\beta s/\epsilon^*\right)\,ds \qquad (9.15)\\
&\quad + V_{\epsilon_n}\left(t_n, X_{t_n}^{\pi,0}, y+\beta\Delta t_n/\epsilon^*\right) - V_{\epsilon_n}(t_n,x,y)\bigg]\\
&\le \Delta t_n e^{-\delta t} U\left(y+\beta\Delta t_n/\epsilon^*\right) + \mathrm{E}\left[\omega\left(|X_{t_n}^{\pi,0}-x|, |\beta\Delta t_n/\epsilon^*|\right)\right]\\
&\to 0
\end{aligned}
$$

when $n\to\infty$, where $\Delta t_n = t_n - t$. The function $\omega : \overline{\mathscr{D}} \to \mathbb{R}$ is a modulus of continuity for V_{ϵ_n} in x and y on $\overline{\mathscr{D}}$, and we see from Theorem 9.14 that a such function ω independent of t and n exists. Equation (9.15) is a contradiction to (9.14), and we see that $\{V_\epsilon\}_{\epsilon>0}$ is lower semi-continuous from the right, uniformly in ϵ. □

Now we will prove left-equicontinuity of $\{V_\epsilon\}_{\epsilon>0}$. Note that the theorem is not valid for $t=T$. The difference between $V_\epsilon(t,x,y)$ and $V_\epsilon(T,x,y)$ will not converge to 0 uniformly in ϵ as $t\to T$: For t close to T and $\epsilon \ll 1$,

$$V_\epsilon(t,x,y) \approx \max_{c\in[0,x]} W(x-c, y+\beta c).$$

On the other hand, we have $V_\epsilon(T,x,y) = W(x,y)$ for all $\epsilon>0$ and $(x,y)\in\overline{\mathscr{D}}$. As we will see in Chapter 10 a consequence of this is that $\{V_\epsilon\}_{\epsilon>0}$ is not uniformly convergent to V on compact subsets of $\overline{\mathscr{D}_T}$, only on compact subsets of $[0,T)\times\overline{\mathscr{D}}$.

If we replace W by $\widehat{W}$ as described in Section 4.1, however, we are been able to prove left-equicontinuity on the whole domain $\overline{\mathscr{D}_T}$. This makes the proof of convergence in Section 10.3 slightly easier, see Section 10.4.

Lemma 9.17. *The family of functions $\{V_\epsilon\}_{\epsilon>0}$ is left-equicontinuous in t on any compact subset $\mathscr{O}_T \subset [0,T)\times\overline{\mathscr{D}}$, i.e., for each $(t,x,y)\in\mathscr{O}_T$ there is an $\omega : [0,T]\to\mathbb{R}$ such that $\omega(0)=0$, ω is continuous at 0 and*

$$|V_\epsilon(t,x,y) - V_\epsilon(t',x,y)| \le \omega(|t-t'|)$$

for all $t'\in[0,t)$ and $\epsilon>0$.

Proof. First we will show that $\{V_\epsilon\}_{\epsilon>0}$ is lower semi-continuous from the left, uniformly in ϵ, i.e., given any $\epsilon'>0$ and $(t,X)\in\mathscr{O}_T$,

$$V_\epsilon(t,X) - V_\epsilon(t',X) < \epsilon'$$

for all $\epsilon>0$ and all $t'\in[0,t)$ sufficiently close to t. By the dynamic programming principle, we have

$$
\begin{aligned}
&V_\epsilon(t,X) - V_\epsilon(t',X)\\
&\quad = V_\epsilon(t,X) - \sup_{(\pi,C)\in\mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_{t'}^{t} e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + V_\epsilon\left(t, X_t^{\pi,C}, Y_t^{\pi,C}\right)\right]\\
&\quad \le V_\epsilon(t,X) - \mathrm{E}\left[\int_{t'}^{t} e^{-\delta s} U\left(Y_s^{0,0}\right) ds + V_\epsilon\left(t, X_t^{0,0}, Y_t^{0,0}\right)\right]\\
&\quad = V_\epsilon(t,X) - \int_{t'}^{t} e^{-\delta s} U\left(ye^{-\beta(s-t')}\right) ds - V_\epsilon\left(t, xe^{\hat{r}(t-t')}, ye^{-\beta(t-t')}\right)\\
&\quad \to 0
\end{aligned}
$$

as $t' \to t$, and the convergence to 0 is uniform is ϵ by Theorem 9.14. We can conclude that $\{V_\epsilon\}_{\epsilon>0}$ is lower semi-continuous from the left, uniformly in ϵ.

Now we want to prove that $\{V_\epsilon\}_{\epsilon>0}$ is upper semi-continuous from the left, uniformly in ϵ, i.e., we want to show that, given any $\epsilon' > 0$ and sequence $\{t_n\}_{n\in\mathbb{N}}$ satisfying $t_n \to t^-$, there is an $N \in \mathbb{N}$ such that

$$
V_\epsilon(t_n, X) - V_\epsilon(t,X) < \epsilon' \tag{9.16}
$$

for all $\epsilon > 0$ and $n \ge N$. To prove equicontinuity, we need to prove that N is independent of ϵ. By the dynamic programming principle, we have

$$
V_\epsilon(t_n,x,y) = \sup_{(\pi_n,C_n)\in\mathscr{B}^\epsilon_{t_n,x,y}} \mathrm{E}\left[\int_{t_n}^{t} e^{-\delta s} U\left(Y_s^{n,\pi_n,C_n}\right) ds + V_\epsilon\left(t, X_t^{n,\pi_n,C_n}, Y_t^{n,\pi_n,C_n}\right)\right], \tag{9.17}
$$

where X_t^{n,π_n,C_n} and Y_t^{n,π_n,C_n} are processes with initial values x and y, respectively, at time t_n. The integral term on the right-hand side of (9.17) converges to 0 as $n \to \infty$, see the proof of (4.16) in Lemma 4.11, and the convergence is uniform in ϵ. It is therefore sufficient to prove that

$$
\mathrm{E}\left[V_\epsilon\left(t, X_t^{n,\pi_n,C_n}, Y_t^{n,\pi_n,C_n}\right)\right] - V_\epsilon(t,x,y) < \epsilon' \tag{9.18}
$$

for all $\epsilon > 0$ and $n \ge N$. We have

$$
\begin{aligned}
&\mathrm{E}\Big[V_\epsilon\left(t, X_t^{n,\pi_n,C_n}, Y_t^{n,\pi_n,C_n}\right)\Big] - V_\epsilon(t,x,y)\\
&\quad = \mathrm{E}\left[V_\epsilon\left(t, X_t^{n,\pi_n,C_n}, Y_t^{n,\pi_n,C_n}\right) - V_\epsilon\left(t, x - \Delta C_n, y + \beta\Delta C_n\right)\right]\\
&\qquad + \mathrm{E}\left[V_\epsilon\left(x - \Delta C_n, y + \beta \Delta C_n\right) - V_\epsilon(t,x,y)\right],
\end{aligned}
$$

where $\Delta C_n := (C_n)_t - (C_n)_{t_n^-}$. The first term on the right-hand side of this equation converges to 0. This follows from the proof of (4.17) in Lemma (4.11). We can prove this by a similar argument as in the proof of (9.18), because V_ϵ is uniformly continuous in x and y and $\mathscr{B}^\epsilon_{t_n,x,y} \subset \mathscr{A}_{t_n,x,y}$. Define $\Delta t_n := t - t_n$. By the definition (8.2) of $\mathscr{B}^\epsilon_{t_n,x,y}$, $\Delta C_n \le \Delta t_n/\epsilon$, and therefore (9.16) is proved if we manage to show that

$$
V_\epsilon(t, x - c_n, y + \beta c_n) - V_\epsilon(t,x,y) < \epsilon' \tag{9.19}
$$

for all $c_n \in [0, \min(x, \Delta t_n/\epsilon)]$ and $\epsilon > 0$ for sufficiently large n. By Lemma 9.13 and $c_n \le x$, we see that there is an $\epsilon^* > 0$ such that (9.19) is satisfied for all $\epsilon < \epsilon^*$ and $n \in \mathbb{N}$.

Now we will prove that (9.19) also is satisfied for $\epsilon > \epsilon^*$. We see that $c_n \to 0$ when $n \to \infty$, and the convergence is uniform in ϵ if $\epsilon > \epsilon^*$. By the uniform equicontinuity of V_ϵ in x and y, we see that (9.19) is satisfied for all ϵ for sufficiently large n. □

Now we have proved equicontinuity in X and t, and we are ready to prove the second main result of the section.

Theorem 9.18. *The family of functions $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous, i.e., for each $(t, x, y) \in [0, T) \times \overline{\mathscr{D}}$, there exists a function ω such that $\omega(0, 0, 0) = 0$, $\omega : [0, \infty) \to [0, \infty)$ is continuous at $(0, 0, 0)$ and*

$$|V_\epsilon(t, x, y) - V_\epsilon(t', x', y')| < \omega(|t - t'|, |x - x'|, |y - y'|)$$

for all $(t', x', y') \in [0, T) \times \overline{\mathscr{D}}$ and $\epsilon > 0$.

Proof. It follows from Lemmas 9.16 and 9.17 that $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous in t on $[0, T) \times \overline{\mathscr{D}}$, so for each $(t, x, y) \in [0, T) \times \overline{\mathscr{D}}$, there exists a function $\hat{\omega}$ that is continuous at 0, and satisfies $\hat{\omega}(0) = 0$ and

$$|V_\epsilon(t, x, y) - V_\epsilon(t', x, y)| < \hat{\omega}(|t - t'|)$$

for all $t' \in [0, T)$. Let $\tilde{\omega} : \overline{\mathscr{D}} \to [0, \infty)$ be a function satisfying the properties in Lemma 9.14. Define $\omega : \overline{\mathscr{D}_T} \to [0, \infty)$ by

$$\omega(t, x, y) = \hat{\omega}(t) + \tilde{\omega}(x, y).$$

We see that ω is continuous at $(0, 0, 0)$, that $\omega(0, 0, 0) = 0$, and that

$$\begin{aligned}
&|V_\epsilon(t, x, y) - V_\epsilon(t', x', y')| \\
&\qquad \le |V_\epsilon(t, x, y) - V_\epsilon(t', x, y)| + |V_\epsilon(t', x, y) - V_\epsilon(t', x', y')| \\
&\qquad \le \hat{\omega}(|t - t'|) + \tilde{\omega}(|x - x'|, |y - y'|) \\
&\qquad = \omega(|t - t'|, |x - x'|, |y - y'|),
\end{aligned}$$

so ω satisfies all the wanted properties. $\square$

Chapter 10
Convergence of the penalty approximation

The main purpose of this chapter is to prove that the viscosity solution V_ϵ of (8.3) and (8.4) converges uniformly to the viscosity solution V of (5.11) and (4.2) on compact subsets of $[0,T) \times \overline{\mathscr{D}}$.

In Section 10.1 we give heuristic arguments of why $V_\epsilon \to V$ when $\epsilon \to 0$. In Section 10.2 we prove the result in the case of increasing stock price, and in Sections 10.3-10.5 we prove the result in the general case. The proofs in Sections 10.3-10.5 use the viscosity solution theory developed in previous chapters, while the proof in Section 10.2 only is based on the original optimization problem described in Chapter 3.

The proofs of Sections 10.3 and 10.4 are relatively similar, except that the one in Section 10.4 is slightly easier, as we have replaced the original terminal utility function W by $\widehat{W}$ as described in Section 4.1. The idea is to use the Arzelà-Ascoli theorem to prove that V_ϵ converges uniformly to some function V', and then prove that V' satisfies (5.11) and (4.2) in a viscosity sense. In Section 10.5 we assume we have a strong comparison principle, and prove directly that V_ϵ converges to V by using weak limits, and without using that $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous and monotone in ϵ.

First we will state the main result of the chapter.

Theorem 10.1. *The viscosity solution V_ϵ of* (8.3) *and* (8.4) *converges uniformly to the viscosity solution V of* (5.11) *and* (4.2) *on compact subsets of* $[0,T) \times \overline{\mathscr{D}}$ *when $\epsilon \to 0$.*

In Sections 10.2, 10.4 and 10.5 we manage to prove the slightly stronger result that V_ϵ converges on compact subsets of $\overline{\mathscr{D}_T}$, because we replace W by $\widehat{W}$ in these sections. The stronger result does not hold for general W, see Section 10.4.

We will prove that $\{V_\epsilon\}_{\epsilon>0}$ converges as a *net* when $\epsilon \to 0$, not only that $\{V_{\epsilon_n}\}_{n\in\mathbb{N}}$ converges uniformly for all sequences $\{\epsilon_n\}_{n\in\mathbb{N}}$ converging to 0. We only need the result concerning sequential convergence when constructing a numerical scheme, but we will prove the stronger result, as the stronger result follows immediately from the weaker result. Since V_ϵ is monotone in ϵ, convergence of the sequence $\{V_{\epsilon_n}\}$ implies convergence of the net $\{V_\epsilon\}_{\epsilon>0}$.

10.1 Heuristic arguments for convergence

In this section we will give two heuristic arguments for convergence. None of the arguments are proofs, but they still show useful techniques for deriving or justifying penalty approximations.

The first argument for convergence is based on the derivation in Chapter 8. The problem described in Chapter 8 is a continuous version of the original optimization problem, where the derivative of C is bounded by $1/\epsilon$. When ϵ decreases, we allow larger values of $c(t)$, and in the limit as $\epsilon \to 0$, we allow discontinuous C. It is therefore reasonable to believe that the solution of the continuous problem converges to the solution of the discontinuous problem as $\epsilon \to 0$.

The second argument for convergence is based on studying equation (8.3). Fix $(t, X) \in \mathscr{D}_T$, suppose V_ϵ satisfies (8.3) in a classical sense, and suppose V_ϵ and all its first and second derivatives converge. Let G^ϵ and F^ϵ denote $G(D_X V_\epsilon)$ and $F\big(t, X, D_X V_\epsilon, D_X^2 V_\epsilon, \mathscr{I}^\pi(t, X, V_\epsilon)\big)$, respectively. Note that F^ϵ is *not* equal to the function F_ϵ in Chapter 8. Consider five different cases:

1. $\frac{1}{\epsilon} G^\epsilon \to 0$ as $\epsilon \to 0$. In this case $(V_\epsilon)_t + F^\epsilon \to 0$ by (8.3), so $\max\{(V_\epsilon)_t + F^\epsilon; G^\epsilon\} \to 0$ as $\epsilon \to 0$.
2. $\frac{1}{\epsilon} G^\epsilon \to p$ as $\epsilon \to 0$ for some $p > 0$. In this case $(V_\epsilon)_t + F^\epsilon \to -p$ and $G^\epsilon \to 0$ as $\epsilon \to 0$, so $\max\{(V_\epsilon)_t + F^\epsilon; G^\epsilon\} \to 0$ as $\epsilon \to 0$.
3. $\frac{1}{\epsilon} G^\epsilon \to -p$ as $\epsilon \to 0$ for some $p > 0$. In this case $(V_\epsilon)_t + F^\epsilon \to 0$ and $G^\epsilon \to 0$ as $\epsilon \to 0$, so $\max\{(V_\epsilon)_t + F^\epsilon; G^\epsilon\} \to 0$ as $\epsilon \to 0$.
4. $\frac{1}{\epsilon} G^\epsilon \to \infty$ as $\epsilon \to 0$. In this case $(V_\epsilon)_t + F^\epsilon \to -\infty$ as $\epsilon \to 0$, so V_ϵ does not converge.
5. $\frac{1}{\epsilon} G^\epsilon \to -\infty$ as $\epsilon \to 0$. In this case $(V_\epsilon)_t + F^\epsilon \to 0$ as $\epsilon \to 0$, so $\max\{(V_\epsilon)_t + F^\epsilon; G^\epsilon\} \to 0$ as $\epsilon \to 0$.

We see that $V_\epsilon \to V$ if both (5.11) and (8.3) have unique classical solutions V and V_ϵ for all $\epsilon > 0$, and V_ϵ and all its derivatives converge.

10.2 Convergence to viscosity solution for increasing stock price

In this section we will prove that V_ϵ converges uniformly to V in the case of increasing stock price. The proof does not apply any of the viscosity theory developed for V and V_ϵ, as we will do in the next three sections. We will prove the theorem by only considering the original optimization problems associated with V and V_ϵ. The proof may therefore represent the most natural approach to the problem for a person with no knowledge of viscosity solutions.

We will assume throughout the section that

(I) S_s is monotonically increasing, i.e., all jumps in the Lévy process are positive and $\sigma = 0$, and
(II) The terminal utility function of V_ϵ is given by $\widehat{W}$, i.e., (8.5) holds.

Note that (I) is not very realistic from a financial point of view, but a theoretical analysis of the case can still be interesting. If (I) holds, the agent will invest all her wealth in the risky asset, as the return always is larger for the risky asset, and our problem is reduced to determining the consumption strategy C.

The idea of the proof is to approximate each control in $(\pi, C) \in \mathscr{A}_{t,x,y}$ by a control $(\pi', C') \in \mathscr{B}^{\epsilon}_{t,x,y}$, and show that we can make the difference $|(V - V_\epsilon)(t,x,y)|$ arbitrarily small by choosing a small enough value for $\epsilon > 0$. Remember that $\mathscr{B}^{\epsilon}_{t,x,y}$ is defined as the set of all absolutely continuous controls in $\mathscr{A}_{t,x,y}$ with derivative bounded by $1/\epsilon$. We will see that V_ϵ converges pointwise to V, and since V and V_ϵ are continuous and V_ϵ is monotone in ϵ, it will follow that the convergence is uniform on compact subsets. We know by Theorem 9.11 that $V_\epsilon \leq V$, so we need to prove that, for any $\epsilon' > 0$ and $(t, X) \in \overline{\mathscr{D}_T}$, we can get $(V - V_\epsilon)(t, X) < \epsilon'$ for sufficiently small $\epsilon > 0$.

We will approximate (π, C) by (π', C') from below, i.e., $C'_s \leq C_s$ for all $s \in [t, T]$. If C' is absolutely continuous with derivative bounded by $1/\epsilon$, we will have $(\pi', C') \in \mathscr{B}^{\epsilon}_{t,x,y}$ by assumption (I), as (I) implies $X_s^{\pi',C'} \geq X_s^{\pi,C}$, see Lemma 10.2 below. If we do not assume S_s is monotonically increasing, it is harder to find an appropriate control $(\pi', C') \in \mathscr{B}^{\epsilon}_{t,x,y}$. If S_s is not monotonically increasing, $C'_s \leq C_s$ does not imply immediately that $X_s^{\pi',C'} \geq X_s^{\pi,C}$, so we may have $(\pi', C') \notin \mathscr{B}^{\epsilon}_{t,x,y}$, even if C' is absolutely continuous with derivative bounded by $1/\epsilon$. It may be possible to generalize the proof such that it is valid for general Lévy processes, but this will not be attempted in this thesis.

Our first lemma states that $(\pi, C') \in \mathscr{A}_{t,x,y}$ if $(\pi, C) \in \mathscr{A}_{t,x,y}$ and $C'_s \leq C_s$ for all $s \in [t, T]$ and all possible outcomes $\omega \in \Omega$. Note that the control π in (π, C') is defined such that it is identical to the control π in (π, C) for all times $s \in [t, T]$ and all ω.

Lemma 10.2. *Assume $(\pi, C) \in \mathscr{A}_{t,x,y}$ for some $(t, x, y) \in [0, T) \times \overline{\mathscr{D}}$, and that C' is a non-negative, non-decreasing adapted process on $[t, T]$, such that $C'_s \leq C_s$ for all $s \in [t, T]$ and $\omega \in \Omega$. Then $(\pi, C') \in \mathscr{A}_{t,x,y}$ and $X_s^{\pi,C'} \geq X_s^{\pi,C}$. If $C'_s = C_s$, we have $Y_s^{\pi,C'} \geq Y_s^{\pi,C}$.*

Proof. We see that $(\pi, C') \in \mathscr{A}_{t,x,y}$ for all $s \in [t, T]$, because $X_s^{\pi,C'} \geq X_s^{\pi,C} \geq 0$ for all $s \in [t, T]$.

Fix $\omega \in \Omega$. The stock price is strictly increasing by (I), and therefore it will always be advantageous to delay consumption as much as possible; if we consume an amount of wealth c at time t instead of at time $t + \Delta t$, we will not have as much advantage of the increase in stock price on the interval $[t, t + \Delta t]$. The consumption strategy C' represents a delay in consumption compared to strategy C. Therefore, we see that $X_s^{\pi,C'} \geq X_s^{\pi,C} \geq 0$, and the first part of the lemma is proved.

If $C_s = C'_s$, we see from (3.3) that

$$\begin{aligned} Y_s^{\pi,C'} &= ye^{-\beta(s-t)} + \beta e^{-\beta s} \int_t^s e^{-\beta s'} \, dC'_{s'} \\ &\geq ye^{-\beta(s-t)} + \beta e^{-\beta s} \int_t^s e^{-\beta s'} \, dC_{s'} \\ &= Y_s^{\pi,C'}, \end{aligned}$$

and the second part of the lemma is proved. □

Now we will prove that any control $(\pi, C) \in \mathscr{A}_{t,x,y}$ can be approximated by an "almost as good" control $(\pi', C') \in \mathscr{A}_{t,x,y}$, such that C' is bounded and constant on some interval $[T-\varsigma, T]$, $\varsigma>0$. We will define C' such that $C'_s \le C_s$ for all $s\in[t,T]$, and Lemma 10.2 will imply that (π', C') is feasible.

Lemma 10.3. *For all* $(t,x,y)\in[0,T)\times\overline{\mathscr{D}}$, $(\pi,C)\in\mathscr{A}_{t,x,y}$ *and* $\epsilon>0$, *there is a* $(\pi',C')\in\mathscr{A}_{t,x,y}$, *a* $\varsigma>0$ *and a* $K\in\mathbb{R}$ *such that*

(1) (π',C') *satisfies*

$$\begin{aligned} \mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\bigg]& \\ <\mathrm{E}\bigg[\int_t^{T-\varsigma} e^{-\delta s}U\left(Y_s^{\pi',C'}\right)\,ds+\widehat{W}\left(X_T^{\pi',C'},Y_T^{\pi',C'}\right)\bigg]+\epsilon,& \end{aligned} \tag{10.1}$$

(2) C' *is constant on* $[T-\varsigma,T]$ *for all* $\omega\in\Omega$, *and*

(3) $C'_s\le K$ *for all* $s\in[t,T]$ *and* $\omega\in\Omega$.

Proof. Define $\Omega_n\subset\Omega$ by

$$\Omega_n=\left\{\omega\in\Omega\,:\,n\le X_T^{\pi,0}<n+1\right\}$$

for each $n\in\mathbb{N}$. Note that the sets Ω_n are disjoint, and that $\Omega=\Omega_\infty\cup\left(\cup_{n=0}^\infty \Omega_n\right)$, where Ω_∞ is some set of probability 0. Using Theorem 2.23, we get

$$\begin{aligned} &\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\bigg]\\ &\quad=\sum_{n=0}^\infty \mathbb{P}(\Omega_n)\,\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\,\Big|\,\Omega_n\bigg]. \end{aligned}$$

Since the sum on the right-hand side of this equation converges, there is an $N\in\mathbb{N}$ such that

$$\begin{aligned} &\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\bigg]\\ &\quad\le\sum_{n=0}^N \mathbb{P}\left(\Omega_n\right)\,\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\,\Big|\,\Omega_n\bigg]+\frac{1}{2}\epsilon. \end{aligned}$$

Define $K:=N+1$,

$$C''_s:=\min\{C_s;\,K\}$$

and $\pi''_s:=\pi_s$ for all $s\in[t,T]$. If $X_T^{\pi,0}<K$, we must have $C_s\le K$ for all $s\in[t,T]$, which implies that $C_s=C''_s$ for all $s\in[t,T]$. It follows that

$$\begin{aligned} &\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\bigg]\\ &\quad\le\sum_{n=0}^N \mathbb{P}\left(\Omega_n\right)\,\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\,\Big|\,\Omega_n\bigg]+\frac{1}{2}\epsilon\\ &\quad=\mathbb{P}\left(0\le X_T^{\pi,0}<K\right)\mathrm{E}\bigg[\int_t^T e^{-\delta s}U\left(Y_s^{\pi,C}\right)\,ds+\widehat{W}\left(X_T^{\pi,C},Y_T^{\pi,C}\right)\,\Big| \end{aligned}$$

$$
\begin{aligned}
&\qquad 0 \leq X_T^{\pi,0} < K\Bigg] + \frac{1}{2}\epsilon \\
&< \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right)\,ds + \widehat{W}\left(X_T^{\pi,C}, Y_T^{\pi,C}\right) \,\Big|\, 0 \leq X_T^{\pi,0} < K\right] + \frac{1}{2}\epsilon \\
&= \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right) \,\Big|\, 0 \leq X_T^{\pi'',0} < K\right] + \frac{1}{2}\epsilon \\
&\leq \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right)\right] + \frac{1}{2}\epsilon.
\end{aligned}
\tag{10.2}
$$

We know that

$$
\begin{aligned}
&\lim_{\varsigma\to 0} \mathrm{E}\left[\int_t^{T-\varsigma} e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right)\right] + \frac{1}{2}\epsilon \\
&\qquad = \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right)\right] + \frac{1}{2}\epsilon,
\end{aligned}
$$

so there is a $\varsigma > 0$ such that

$$
\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right)\right] \\
&\qquad < \mathrm{E}\left[\int_t^{T-\varsigma} e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right)\right] + \frac{1}{2}\epsilon.
\end{aligned}
\tag{10.3}
$$

Define $\pi' := \pi$, $C'_s := C''_s \mathbf{1}_{s\leq T-\varsigma} + C''_{T-\varsigma}\mathbf{1}_{s>T-\varsigma}$, $\varDelta C := C''_T - C''_{T-\varsigma} = C''_T - C'_T$ and $C'''_s := C'_s + \varDelta C \mathbf{1}_{s=T}$. By applying Lemma 10.2 with C'' and C''', we see that

$$
X_T^{\pi'',C''} \leq X_T^{\pi'',C'''} = X_T^{\pi',C'} - \varDelta C
$$

and

$$
Y_T^{\pi'',C''} \leq Y_T^{\pi'',C'''} = Y_T^{\pi',C'} + \beta \varDelta C.
$$

Using these two inequalities and assumption (II), we get

$$
\widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right) \leq \widehat{W}\left(X_T^{\pi',C'} - \varDelta C, Y_T^{\pi',C'} + \beta \varDelta C\right) \leq \widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right).
$$

We see from (10.3) that

$$
\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi'',C''}, Y_T^{\pi'',C''}\right)\right] \\
&\qquad < \mathrm{E}\left[\int_t^{T-\varsigma} e^{-\delta s} U\left(Y_s^{\pi'',C''}\right)\,ds + \widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right)\right] + \frac{1}{2}\epsilon \\
&\qquad \leq \mathrm{E}\left[\int_t^{T} e^{-\delta s} U\left(Y_s^{\pi',C'}\right)\,ds + \widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right)\right] + \frac{1}{2}\epsilon.
\end{aligned}
\tag{10.4}
$$

Combining (10.2) and (10.4), we get (10.1). □

Now we will prove that the kind of controls described in Lemma 10.3 can be approximated from below by a control in $\mathscr{B}_{t,x,y}^{\epsilon'}$ for some ϵ'. Again Lemma 10.2 will imply feasibility of the new control.

Lemma 10.4. *Let $(t,x,y) \in [0,T) \times \overline{\mathscr{D}}$, $(\pi, C) \in \mathscr{A}_{t,x,y}$ and $\epsilon > 0$, and suppose there is a $\varsigma > 0$ and a $K \in \mathbb{R}$, such that $C_s \leq K$ for all $s \in [t,T]$ and C is constant on $[T-\varsigma, T]$. Then there is an $\epsilon' > 0$ and a $(\pi', C') \in B_{t,x,y}^{\epsilon'}$ such that $\pi' \equiv \pi$,*

(1) $C'_s \leq C_s$ *for all* $s \in [t,T]$ *and* $\omega \in \Omega$*,*
(2) $C_s - C'_s < \epsilon$ *for all* $s \in [t,T]\backslash U_\omega$*, where* $U_\omega \subset [t,T]$ *satisfies* $\lambda(U_\omega) < \epsilon$ *and* λ *is the Lebesgue measure,*
(3) $C_s - C'_s \leq K$ *for all* $s \in [t,T]$*, and*
(4) $C'_T = C_T$*.*

Proof. Define

$$\epsilon' = \frac{1}{2}\min\left\{\frac{\varsigma}{K}; \frac{\epsilon}{K}\right\}, \qquad \pi'_s = \pi_s \quad \text{and} \quad C'_s = \int_t^s c(s')\,ds' \tag{10.5}$$

for all $s \in [t,T]$, where

$$c(s) = \begin{cases} 0 & \text{if } C_s - C'_s \leq 0,\ s < T-\varsigma, \\ \frac{1}{\epsilon'} & \text{if } C_s - C'_s \geq \epsilon,\ s < T-\varsigma, \\ \lim_{r\to s^-} c(r) & \text{if } 0 < C_s - C'_s < \epsilon,\ s < T-\varsigma, \\ \frac{1}{\epsilon'} & \text{if } T-\varsigma \leq s \leq T-\varsigma+\epsilon'(C_{T-\varsigma} - C'_{T-\varsigma}), \\ 0 & \text{if } s > T-\varsigma+\epsilon'(C_{T-\varsigma} - C'_{T-\varsigma}). \end{cases}$$

The limit $\lim_{r\to s^-} c(r)$ is well-defined for $s \in (t, T-\varsigma]$ by the following argument: Suppose $c(s_1) = 1/\epsilon'$ for some $s_1 < T-\varsigma$, and define

$$s_2 = \inf\left\{s \in [t,s_1] : c(r) = 1/\epsilon' \text{ for all } r \in [s,s_1]\right\}. \tag{10.6}$$

We must have $C_{s_2} - C'_{s_2} \geq \epsilon$, as $C_s - C'_s$ is right-continuous, and therefore $c(s_2) = 1/\epsilon'$. Since C_s is increasing, and C'_s is continuous with gradient less than or equal to $1/\epsilon'$, we see that $C_s - C'_s > 0$ on $[s_2, s_2+\epsilon\epsilon')$, and therefore $c(s) = 1/\epsilon'$ on $[s_2, s_2+\epsilon\epsilon']$. We conclude that c does not take the value $1/\epsilon'$ pointwise or on arbitrarily small intervals of $[t, T-\varsigma]$, but on intervals of length at least $\epsilon\epsilon'$. It follows that the value $\lim_{r\to s^-} c(r)$ is well-defined for all $s \in (t,T]$.

Note that C' is an adapted process, as it only uses information up to time s to decide C'_s. We see that

(1) $C'_s \leq C_s$ for all $s \in [t,T]$ and $\omega \in \Omega$,
(2) $C_s - C'_s \geq \epsilon$ only when $s \in U_\omega := \{s \in [t,T] : c(s) = 1/\epsilon'\}$. The process C'_s is monotonically increasing, $C'_0 = 0$ and $C'_T \leq K$, and therefore $\lambda(U_\omega) \leq K\epsilon' < \epsilon$,
(3) $C_s - C'_s \leq K$ for all $s \in [t,T]$, because $C' \geq 0$ and $C \leq K$ everywhere, and
(4) We have $c(s) = 1/\epsilon'$ for $s \in \left[T-\varsigma, T-\varsigma+\epsilon'(C_{T-\varsigma} - C'_{T-\varsigma})\right]$. Therefore $C'_s = C_s$ for $s = T-\varsigma+\epsilon'(C_{T-\varsigma} - C'_{T-\varsigma})$, and we see that $C'_s = C_s$ for all $s > T-\varsigma+\epsilon'(C_{T-\varsigma} - C'_{T-\varsigma})$. We see that $T-\varsigma+\epsilon'(C_{T-\varsigma} - C'_{T-\varsigma}) < T$ by the definition of ϵ'.

It follows from (1) and Lemma 10.2 that $(\pi', C') \in B^{\epsilon'}_{t,x,y}$. □

Now we will prove that the approximation $(\pi', C') \in \mathscr{B}^{\epsilon'}_{y,x,y}$ of $(\pi, C) \in \mathscr{A}_{t,x,y}$ defined in Lemma 10.4, gives values for $X^{\pi',C'}$ and $Y^{\pi',C'}$ that are close to or larger than the "original" values $X^{\pi,C}$ and $Y^{\pi,C}$.

Lemma 10.5. *Suppose* $(t,x,y) \in [0,T) \times \overline{\mathscr{D}}$*,* $(\pi, C) \in \mathscr{A}_{t,x,y}$ *and* $\epsilon > 0$*, where* $C \leq K$ *everywhere for some* $K \in \mathbb{R}$*, and* C *is constant on* $[T-\varsigma, T]$*. Then there is an* $\epsilon' > 0$ *and a* $(\pi', C') \in \mathscr{B}^{\epsilon'}_{t,x,y}$ *such that*

(1) $X_s^{\pi',C'} \geq X_s^{\pi,C}$ *for all* $s \in [t,T]$,
(2) $Y_T^{\pi',C'} \geq Y_T^{\pi,C}$,
(3) $Y_s^{\pi,C} - Y_s^{\pi',C'} < \epsilon$ *for all* $s \in [t,T]\backslash U_\omega$, *where* $U_\omega \subset [t,T]$ *is such that* $\lambda\big([t,T]\backslash U_\omega\big) < \epsilon$, *and*
(4) $Y_s^{\pi,C} - Y_s^{\pi',C'} < K$ *for all* $s \in [t,T]$.

Proof. By Lemma 10.4, there exists an $\epsilon' > 0$, a $K' > 0$ and a $(\pi', C') \in \mathscr{B}_{t,x,y}^{\epsilon'}$ such that

(1)' $C'_s \leq C_s$ for all $s \in [t,T]$ and $\omega \in \Omega$,
(2)' $C_s - C'_s < \epsilon''$ for all $s \in [t,T]\backslash U_\omega$, where $\omega \in \Omega$ and $U_\omega \subset [t,T]$ satisfies $\lambda\big(U_\omega\big) < \epsilon''$,
(3)' $C_s - C'_s \leq K'$ for all $s \in [t,T]$, and
(4)' $C'_T = C_T$,

where
$$\epsilon'' = \min\left\{\frac{\epsilon}{4\beta(T-t)}; \frac{\epsilon}{4\beta K'}\right\}. \tag{10.7}$$
The claims (1) and (2) of our lemma follow directly from Lemma 10.2. We see easily that (3) also is valid, because

$$\begin{aligned}
Y_s^{\pi,C} - Y_s^{\pi',C'} &= \beta e^{-\beta s}\int_t^s e^{\beta s'} d(C_{s'} - C'_{s'})\\
&\leq \beta e^{-\beta s}\left[(s-t)e^{\beta s}\epsilon'' + \lambda\big(U_\omega\big)e^{\beta s}K'\right]\\
&< \epsilon'
\end{aligned}$$

for all $s \in [t,T]\backslash U_\omega$. By defining $K := \beta(s-t)K$, (4) also holds, because
$$Y_s^{\pi,C} - Y_s^{\pi',C'} \leq \beta e^{-\beta s}(s-t)e^{\beta s}K' = K.$$

□

The lemma below proves that any $(\pi, C) \in \mathscr{A}_{t,x,y}$ can be approximated by a $(\pi', C') \in \mathscr{B}_{t,x,y}^{\epsilon'}$ that gives almost as good utility for the agent as (π, C). First we apply Lemma 10.3 to find an approximation to (π, C) that is bounded and constant close to T, and then we apply Lemma 10.5 to find an approximation in $\mathscr{B}_{t,x,y}^{\epsilon'}$ for some $\epsilon' > 0$.

Lemma 10.6. *For all* $(t,x,y) \in [0,T) \times \overline{\mathscr{D}}$, $(\pi, C) \in \mathscr{A}_{t,x,y}$ *and* $\epsilon > 0$, *there is an* $\epsilon' > 0$ *and a* $(\pi', C') \in \mathscr{B}_{t,x,y}^{\epsilon'}$ *such that*

$$\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + \widehat{W}\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right]\\
&\qquad \leq \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi',C'}\right) ds + \widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right)\right] + \epsilon.
\end{aligned} \tag{10.8}$$

Proof. By (u_{ii}) in Chapter 3, there exists a $K_\epsilon > 0$ such that
$$U\left(|y_1 - y_2|\right) \leq \frac{\epsilon}{4(T-t)} + K_\epsilon |y_1 - y_2|^\gamma \tag{10.9}$$
for all $y_1, y_2 \in [0, \infty)$. By Lemma 10.5, there exists an $\epsilon' > 0$ and a $(\pi', C') \in \mathscr{B}_{t,x,y}^{\epsilon'}$ such that

(1) $X_s'^{\pi',C'} \geq X_s^{\pi,C}$ for all $s \in [t,T]$,
(2) $Y_T'^{\pi',C'} \geq Y_T^{\pi,C}$,
(3) $Y_s^{\pi,C} - Y_s'^{\pi',C'} < \epsilon''$ for all $s \in [t,T]\backslash U_\omega$, where $U_\omega \subset [t,T]$ is such that $\lambda([t,T]\backslash U_\omega) < \epsilon''$, and
(4) $Y_s^{\pi,C} - Y_s'^{\pi',C'} < K$ for all $s \in [t,T]$,

where

$$\epsilon'' < \min\left\{\left(\frac{\epsilon}{4|T-t|K_\epsilon}\right)^{1/\gamma}; \frac{\epsilon}{4K_\epsilon K^\gamma}; |T-t|\right\}. \tag{10.10}$$

By (1) and (2), we have

$$\mathrm{E}\left[\widehat{W}\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right] \leq \mathrm{E}\left[\widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right)\right]. \tag{10.11}$$

By (10.9), (3) and (4), we have

$$\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds\right] - E\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi',C'}\right) ds\right] \\
&\quad \leq |T-t|\left(\frac{\epsilon}{4(T-t)} + K_\epsilon(\epsilon'')^\gamma\right) + \lambda\left(U_\omega\right)\left(\frac{\epsilon}{4(T-t)} + K_\epsilon K^\gamma\right) \\
&\quad < \epsilon.
\end{aligned} \tag{10.12}$$

Combining (10.11) and (10.12), we get (10.8). □

Using the above lemma and the monotonicity of V_ϵ in ϵ, we see that V_ϵ converges pointwise to V.

Lemma 10.7. *The family of functions $\{V_\epsilon\}_{\epsilon>0}$ converges pointwise to V as $\epsilon \to 0$.*

Proof. We see that V_ϵ is pointwise convergent, since $V_\epsilon(t,x,y)$ is bounded and monotone in ϵ by Theorem 9.11. Define $V' : [0,T) \times \overline{\mathscr{D}} \to \mathbb{R}$ by

$$V'(t,x,y) = \lim_{\epsilon \to 0} V_\epsilon(t,x,y).$$

We want to show that $V' = V$. We see immediately that the result holds for $t = T$, as $V_\epsilon = V$ for $t = T$ by assumption (II). We will therefore concentrate on the case $t < T$ from now on.

$V' \leq V$, because $V_\epsilon \leq V$ for all $\epsilon > 0$ by Theorem 9.11. It is therefore sufficient to show that $V' \geq V$. Let $(\pi, C) \in \mathscr{A}_{t,x,y}$ satisfy

$$V(t,x,y) = E\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + \widehat{W}\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right].$$

The existence of an optimal control follows from Theorem 3.6. By Lemma 10.6, there exist an $\epsilon' > 0$ and a $(\pi', C') \in \mathscr{B}_{t,x,y}^{\epsilon'}$ such that

$$\begin{aligned}
&\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + \widehat{W}\left(X_T^{\pi,C}, Y_T^{\pi,C}\right)\right] \\
&\quad \leq \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi',C'}\right) ds + \widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right)\right] + \epsilon.
\end{aligned}$$

We have

$$
\begin{aligned}
V(t,x,y) \leq & \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi',C'}\right) ds + \widehat{W}\left(X_T^{\pi',C'}, Y_T^{\pi',C'}\right)\right] + \epsilon \\
\leq & V_{\epsilon'}(t,x,y) + \epsilon \\
\leq & V'(t,x,y) + \epsilon.
\end{aligned}
$$

Letting $\epsilon \to 0$, we see that $V' \geq V$, so $V = V'$, and the lemma is proved. $\square$

What remains to show, is that the convergence of V_ϵ to V is uniform on compact subsets of $\overline{\mathscr{D}_T}$. This follows directly from the fact that V_ϵ is monotone in ϵ, and that V_ϵ and V are continuous.

Proof of Theorem 10.1 in the case of increasing stock price: We will show by contradiction that V_ϵ converges uniformly to V on compact subsets of $\overline{\mathscr{D}_T}$. Suppose there is an $\epsilon > 0$ and sequences $\{\epsilon_n\}_{n\in\mathbb{N}}$ and $\{(t_n, X_n)\}_{n\in\mathbb{N}}$, such that

$$
|V(t_n, X_n) - V_{\epsilon_n}(t_n, X_n)| > \epsilon,
$$

where $\lim_{n\to\infty} \epsilon_n = 0$ and (t_n, X_n) is in some compact subset of $\overline{\mathscr{D}_T}$ for all $n \in \mathbb{N}$. We can assume without loss of generality that $\{\epsilon_n\}$ is decreasing, and that $(t_n, X_n) \to (t, X)$ as $n \to \infty$ for some $(t, X) \in \overline{\mathscr{D}_T}$. Since $V_\epsilon \leq V$ for all $\epsilon > 0$, we know that

$$
V(t_n, X_n) - V_{\epsilon_n}(t_n, X_n) > \epsilon \tag{10.13}
$$

for all $n \in \mathbb{N}$. Pick an $N \in \mathbb{N}$ such that

$$
V(t, X) - V_{\epsilon_N}(t, X) < \epsilon/2. \tag{10.14}
$$

Since V_{ϵ_N} and V are continuous and $(t_n, X_n) \to (t, X)$, there is an $M \in \mathbb{N}$, $M \geq N$, such that

$$
|V_{\epsilon_N}(t_M, X_M) - V_{\epsilon_N}(t, X)|, \; |V(t_M, X_M) - V(t, X)| < \epsilon/4. \tag{10.15}
$$

Using (10.14), (10.15) and Theorem 9.11, we get

$$
\begin{aligned}
V(t_M, X_M) - V_{\epsilon_M}(t_M, X_M) < & V(t_M, X_M) - V_{\epsilon_N}(t_M, X_M) \\
< & \big(V(t, X) + \epsilon/4\big) - \big(V_{\epsilon_n}(t, X) - \epsilon/4\big) \\
= & V(t, X) - V_{\epsilon_N}(t, X) + \epsilon/2 \\
< & \epsilon,
\end{aligned}
$$

which is a contradiction to (10.13). It follows that the convergence of V_ϵ to V is uniform on compact subsets of $\overline{\mathscr{D}_T}$, and the theorem is proved. $\square$

When we construct a numerical scheme in Part III, we are interested in the convergence rate of V_ϵ to V. We can derive an approximate worst-case convergence rate by following the argument in this section. Given an $\epsilon > 0$ and a $(t, X) \in [0, T) \times \overline{\mathscr{D}}$, we want to find an $\epsilon' > 0$ such that $|V(t, X) - V_{\epsilon'}(t, X)| < \epsilon$.

Based on, respectively, equations (10.10), (10.7) and (10.5) in Lemmas 10.6, 10.5 and 10.4, we define ϵ'', ϵ''' and $\epsilon' > 0$ as follows:

$$\begin{aligned}
\epsilon'' &= \min\left\{\left(\frac{\epsilon}{4|T-t|}\right)^{1/\gamma}; \frac{\epsilon}{4K_\epsilon^\gamma}; |T-t|\right\}\\
\epsilon''' &= \frac{1}{2}\min\left\{\frac{\epsilon''}{4\beta(T-t)}; \frac{\epsilon''}{4\beta K_{\epsilon''}}\right\}\\
\epsilon' &= \frac{1}{2}\min\left\{\frac{\varsigma_{\epsilon'''}}{K_{\epsilon'''}}; \frac{\epsilon'''}{K_{\epsilon'''}}\right\}.
\end{aligned} \tag{10.16}$$

We define $K_{\hat\epsilon}$ and $\varsigma_{\hat\epsilon}$ for $\hat\epsilon = \epsilon, \epsilon'', \epsilon'''$ to be the constants mentioned in Lemmas 10.5, 10.4 and 10.3, respectively, when ϵ is replaced by $\hat\epsilon$. The constant K_ϵ in (10.10) is independent of ϵ if U is a CRRA utility function, and is therefore assumed to be constant and not included in (10.16). By following the arguments in the proofs above, we see that $|V(t,X) - V_{\epsilon'}(t,X)| < \epsilon$ if (10.16) holds.

We see from (10.3) that $\varsigma_{\hat\epsilon}$ is of order $\hat\epsilon$. We see from the proof of Lemma 10.3 that $K_{\hat\epsilon}$ increases as $\hat\epsilon \to 0$, but the increase rate of $K_{\hat\epsilon}$ is low, as $\mathbb{P}\left(X_T^{\pi,0} > K_{\hat\epsilon}\right)$ converges fast to 0 when $K_{\hat\epsilon}$ increases. We will therefore assume $K_{\hat\epsilon}$ is independent of $\hat\epsilon$.

We see from (10.16) that $\epsilon'' \approx A_1\epsilon^{1/\gamma}$, $\epsilon''' \approx A_2\epsilon''$ and $\epsilon' \approx A_3\epsilon'''$ for constants $A_1, A_2, A_3 > 0$ when ϵ is small, so $\epsilon' \approx A_1A_2A_3\epsilon^{1/\gamma}$. It follows that γ is an approximate worst-case convergence rate.

10.3 Convergence to viscosity solution, proof I

This section contains our first proof of Theorem 10.1 in the general case, and it is the only section where we do not assume the terminal condition of V_ϵ is given by $\widehat{W}$. The proof consists of several steps. First we will prove that V_ϵ converges pointwise to a function V' on $[0,T)\times\overline{\mathscr{D}}$, and, by applying the Arzelà-Ascoli theorem and equicontinuity results from Section 9.2, we will prove that V' is continuous and that the convergence is uniform on compact sets. Then we prove that $V' \to V$ as $t \to T$, and that V' is a viscosity supersolution. We prove that V' is a viscosity supersolution by using that V_ϵ is a viscosity supersolution of the penalized problem. By applying the comparison principle 6.15 and Theorem 9.11, we see that $V' = V$.

Note that it is not necessary to prove the viscosity subsolution property of V'. Instead of proving the viscosity subsolution property of V', we apply the fact that $V_\epsilon \le V$ for all $\epsilon > 0$. The inequality $V_\epsilon \le V$ implies that $V' \le V$. The comparison principle and the viscosity supersolution property of V' implies that $V' \ge V$. It follows that $V' = V$, and we manage to prove this without proving that V' is a viscosity subsolution.

First we prove, using Theorem 9.11, that V_ϵ converges pointwise.

Lemma 10.8. *The family of functions* $\{V_\epsilon\}_{\epsilon>0}$ *converges pointwise to a function* V' *on* $[0,T)\times\overline{\mathscr{D}}$ *when* $\epsilon \to 0$.

Proof. Fix (t,X), and consider a decreasing sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ converging towards 0. By Theorem 9.11, $\{V_{\epsilon_n}(t,X)\}_{n\in\mathbb{N}}$ is a monotonically increasing sequence that is bounded by $V(t,X)$, and therefore it is convergent. As described in the introduction to this chapter, $\{V_\epsilon(t,X)\}_{\epsilon>0}$ converges as a net, not only sequentially, since V_ϵ is monotone in ϵ. □

Based on the lemma above, we can define a function V' by the following definition.

Definition 10.9. *Define* $V' : [0,T) \times \overline{\mathscr{D}} \to \mathbb{R}$ *by*

$$V'(t,x,y) := \lim_{\epsilon \to 0} V_\epsilon(t,x,y).$$

Note that V' is not (yet) defined for $t = T$. We will extend the domain of V' to include $t = T$ later.

Lemma 10.8 says that V_ϵ converges pointwise. However, it does not say anything about the nature of the convergence, or of properties of the function V'. We will focus on these properties in the next theorems. First we apply the Arzelà-Ascoli theorem to prove that all sequences $\{V_{\epsilon_n}\}_{n\in\mathbb{N}}$ have a subsequence that converges uniformly to V'. We will also see that V' is continuous.

Lemma 10.10. *For any sequence* $\{\epsilon_n\}_{n\in\mathbb{N}}$ *converging towards 0,* $\{V_{\epsilon_n}\}_{n\in\mathbb{N}}$ *has a subsequence that converges uniformly to* V' *on compact subsets of* $[0,T) \times \overline{\mathscr{D}}$*.*

Proof. V_ϵ satisfies the conditions of the Arzelà-Ascoli theorem (Theorem 2.25) by Theorems 9.3, 9.2 and 9.18. By the Arzelà-Ascoli theorem, we see that $\{V_{\epsilon_n}\}_{n\in\mathbb{N}}$ has a subsequence that converges uniformly to a function V''. We know that $\{V_{\epsilon_n}\}_{n\in\mathbb{N}}$ converges pointwise to V', and this implies that $V'' = V'$. □

Corollary 10.11. V' *is continuous.*

Proof. This follows directly from Lemma 10.10, as the limit function in the Arzelà-Ascoli theorem is continuous. □

We will use Lemma 10.10 to show that $\{V_\epsilon\}_{\epsilon>0}$ converges uniformly to V' as a net. First we need to prove the following technical lemma.

Lemma 10.12. *Let* $f, f_n : A \to \mathbb{R}$ *for some* $A \subseteq \mathbb{R}^m$ *for all* $n \in \mathbb{N}$*. Suppose* $\{f_n\}_{n\in\mathbb{N}}$ *is such that every subsequence* $\{f_{n_j}\}_{j\geq 1}$ *has a further subsequence* $\{f_{n_{j_k}}\}_{k\geq 1}$ *that converges to* f *in the* L_∞ *norm. Then the sequence* $\{f_n\}_{n\in\mathbb{N}}$ *converges to* f *in the* L_∞ *norm.*

Proof. Suppose $\{f_n\}_{n\in\mathbb{N}}$ does not converge to f. Then there exist an $\epsilon > 0$ and a sequence $\{f_{n_j}\}_{j\geq 1}$ such that $\|f_{n_j} - f\|_{L_\infty} > \epsilon$ for $j = 1, 2, \ldots$. The sequence $\{f_{n_j}\}_{j\geq 1}$ has no subsequence $\{f_{n_{j_k}}\}_{k\geq 1}$ that converges to f. We have obtained a contradiction, and therefore we see that $\{f_n\}_{n\in\mathbb{N}}$ converges to f. □

We are now ready to prove our first main result.

Theorem 10.13. *The family of functions* $\{V_\epsilon\}_{\epsilon>0}$ *converges uniformly to* V' *on all compact subsets of* $[0,T) \times \overline{\mathscr{D}}$*, and* V' *is continuous.*

Proof. We have already proved that V' is continuous (Corollary 10.11), so we will focus on proving the first part of the theorem. Let $\{\epsilon_n\}_{n\in\mathbb{N}}$ be any sequence converging to 0. As described in the introduction to this chapter, it is sufficient to show that $\{V_{\epsilon_n}\}_{n\in\mathbb{N}}$ converges uniformly to V' on any compact subsets of $[0,T) \times \overline{\mathscr{D}}$. By Lemma 10.10, we know that every subsequence of $\{V_{\epsilon_n}\}$ has a further subsequence that converges uniformly to a continuous function V''. Since V_ϵ converges to V', we see that $V'' = V'$. For continuous functions, convergence in the L_∞ norm is the same as uniform convergence. Therefore we can use Lemma 10.12 to deduce that $\{V_{\epsilon_n}\}$ itself converges uniformly to V'. □

We will now prove that V' is a viscosity solution of (5.11) and (4.2). By uniqueness results, it will follow that $V' = V$, where V is the value function of the original problem. First we need to extend the domain of V' to also contain points where $t = T$. We want to extend V' in such a way that V' becomes continuous, as this is one condition that must be satisfied in order for V' to be a viscosity solution of (5.11) and (4.2).

Lemma 10.14. *The function V' satisfies $V'(t,x,y) \to V(T,x,y)$ as $t \to T$.*

Proof. The statement will be proved by contradiction. If $V'(t,x,y) \nrightarrow V(T,x,y)$, at least one of the following statements must be true:

(1) There exist an $\epsilon' > 0$ and a sequence $\{t_n\}_{n\in\mathbb{N}}$ such that $t_n \to T^-$ and $V'(t_n,x,y) - V(T,x,y) > \epsilon'$ for all $n \in \mathbb{N}$.
(2) There exist an $\epsilon' > 0$ and a sequence $\{t_n\}_{n\in\mathbb{N}}$ such that $t_n \to T^-$ and $V(T,x,y) - V'(t_n,x,y) > \epsilon'$ for all $n \in \mathbb{N}$.

First we consider (1). Since V is continuous at (T,x,y) and $t_n \to T$, $V(t_n,x,y) < V(T,x,y) + \epsilon'$ for some $n \in \mathbb{N}$. By Theorem 9.11, we also know that $V'(t_n,x,y) < V(t_n,x,y)$. Combining these inequalities we get $V'(t_n,x,y) < V(T,x,y) + \epsilon'$, which is a contradiction to the assumption we made in (1).

Now we consider (2). Find a $t_{\epsilon'} \in [0,T)$ and a function $\hat{\epsilon} : [t_{\epsilon'},T) \to \mathbb{R}$ such that

$$V(T,x,y) - V_{\hat{\epsilon}(t)}(t,x,y) < \epsilon' \tag{10.17}$$

for all $t \in [t_{\epsilon'},T)$. First we need to show that a such construction is possible.

Define $t_{\epsilon'}$ to be such that

$$W\left(X_T^{0,0}, Y_T^{0,0}\right) > W(x-c^*, y+\beta c^*) - \epsilon'/2$$

for all $t > t_{\epsilon'}$, where $X_s^{0,0}$ and $Y_s^{0,0}$ are processes taking initial value $x - c^*$ and $y + \beta c^*$, respectively, at time t, and $c^* \in [0,x]$ is chosen such that $W(x-c^*, y+\beta c^*)$ is maximized. We know that a $t_{\epsilon'}$ with the wanted properties can be found, because the development of $X_s^{\pi,C}$ and $Y_s^{\pi,C}$ is deterministic and continuous if $(\pi,C) = (0,0)$. We also know that

$$\begin{aligned} V_\epsilon(t, x-c^*, y+\beta c^*) &= \sup_{(\pi,C)\in\mathscr{B}^\epsilon_{t,x,y}} \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{\pi,C}\right) ds + W(X_T^{\pi,C}, Y_T^{\pi,C})\right] \\ &\geq \mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{0,0}\right) ds + W\left(X_T^{0,0}, Y_T^{0,0}\right)\right] \\ &\geq W\left(X_T^{0,0}, Y_T^{0,0}\right), \end{aligned}$$

so

$$V_\epsilon(t, x-c^*, y+\beta c^*) > W(x-c^*, y+\beta c^*) - \epsilon'/2$$

for all $\epsilon > 0$ and $t \in [t_{\epsilon'},T)$.

Now fix $t \in [t_{\epsilon'},T)$, and define $\hat{\epsilon}(t)$ to be such that

$$V_{\hat{\epsilon}(t)}(t,x,y) > \max_{c\in[0,x]} V_{\hat{\epsilon}(t)}(t,x-c,y+\beta c) - \epsilon'/2.$$

It is possible to define such an $\hat{\epsilon}(t)$ by Lemma 9.13. We see that (10.17) is satisfied, as

$$\begin{aligned}
V(T,x,y) &- V_{\hat{\epsilon}(t)}(t,x,y) \\
&< W(x-c^*,y+\beta c^*) - \max_{c\in[0,x]} V_{\hat{\epsilon}(t)}(t,x-c,y+\beta c) + \epsilon'/2 \\
&< \big(V_{\hat{\epsilon}(t)}(t,x-c^*,y+\beta c^*) + \epsilon'/2\big) - \max_{c\in[0,x]} V_{\hat{\epsilon}(t)}(t,x-c,y+\beta c) + \epsilon'/2 \\
&\leq \epsilon'
\end{aligned}$$

for all $t \in [t_{\epsilon'}, T]$. Now choose $n \in \mathbb{N}$ such that $t_n \geq t_{\epsilon'}$. We have

$$V(T,x,y) - V_{\hat{\epsilon}(t_n)}(t_n,x,y) < \epsilon'.$$

This is a contradiction to (2), since $V_{\hat{\epsilon}(t_n)}(t_n,x,y) < V'(t_n,x,y)$ by Theorem 9.11. □

We extend the domain of V' to $t = T$ by defining

$$V'(T,x,y) = \lim_{t\to T^-} V'(t,x,y) = V(T,x,y).$$

By Lemma 10.14 and Theorem 10.13, we see that V' is continuous. Now as V' is defined on the whole domain $\overline{\mathscr{D}_T}$, we can prove that it is a viscosity supersolution of (5.11) and (4.2).

Lemma 10.15. *The function V' is a viscosity supersolution of* (5.11) *and* (4.2) *on* $\mathscr{D}_T$.

Proof. We have $V'(T,X) = V(T,X)$ for all $X \in \overline{\mathscr{D}}$ by definition, and therefore V' satisfies (4.2). By Theorem 10.13 and Lemma 10.14, we see that V' is continuous. What is left to prove, is that V' is a viscosity supersolution of (5.11) on $\mathscr{D}_T$.

Assume $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t,X) \in \mathscr{D}_T$ satisfy

(1) (t,X) is a global minimum of $V' - \phi$ on $\mathscr{D}_T$,
(2) $\phi(X) > V'(X)$,
(3) ϕ has compact support, and
(4) there is a constant $\epsilon' > 0$ such that $V'(t,X) - \phi(t,X) < V'(t',X') - \phi(t',X') - \epsilon'$ for all other minima (t',X') of $V' - \phi$.

We see that V' is continuous by Theorems 10.14 and 10.13, and therefore V' a viscosity supersolution of (5.11) if we can prove that:

$$\max\big\{G(D_X\phi);\ \phi_t + F\big(t,X,D_X\phi,D_X^2\phi,\mathscr{I}^\pi(t,X,\phi)\big)\big\} \leq 0. \tag{10.18}$$

Assume $\{\epsilon_n\}_{n\in\mathbb{N}}$ is a sequence of positive numbers converging to 0, and that $(t_{\epsilon_n}, X_{\epsilon_n})$ is a global minimum of $V_{\epsilon_n} - \phi$. First we will show that $(t_{\epsilon_n}, X_{\epsilon_n}) \to (t,X)$ as $n \to \infty$. Fix $\epsilon_a > 0$. We want to show that $|(t_{\epsilon_n}, X_{\epsilon_n}) \to (t,X)| < \epsilon_a$ for all sufficiently large n.

By assumption (4), there is a ball B around (t,X) and an $\epsilon_b > 0$, such that

$$(V' - \phi)(t,X) < (V' - \phi)(t',X') - \epsilon_b$$

for all $(t',X') \in \overline{\mathscr{D}_T} \backslash B$. We may choose B so small that it has radius smaller than ϵ_a.

By Theorem 10.13, V_ϵ converges uniformly to V' on $\mathrm{supp}(\phi)$. Therefore there exists an $\epsilon_c > 0$ such that $V'(t',X') - V_\epsilon(t',X') < \epsilon_b$ for all $(t',X') \in \mathrm{supp}(\phi)$ and all $\epsilon < \epsilon_c$.

All global minima of $V_\epsilon - \phi$ must be inside B for $\epsilon < \epsilon_c$:

1. For $(t', X') \in \text{supp}(\phi)\backslash B$, we have
$$\begin{aligned}(V_\epsilon - \phi)(t', X') &= (V_\epsilon - V')(t', X') + (V' - \phi)(t', X')\\ &> (-\epsilon_b) + \big((V' - \phi)(t, X) + \epsilon_b\big)\\ &= (V' - \phi)(t, X)\\ &\geq (V_\epsilon - V')(t, X) + (V' - \phi)(t, X)\\ &= (V_\epsilon - \phi)(t, X),\end{aligned}$$
so a global minimum for $V_\epsilon - \phi$ cannot lie in $\text{supp}(\phi)\backslash B$.
2. For $(t', X') \in \overline{\mathscr{D}_T}\backslash\text{supp}(\phi)$, we have $(V_\epsilon - \phi)(t', X') = V_\epsilon(t', X') \geq 0$. We also know that $(V_\epsilon - \phi)(t, X) = (V_\epsilon - V')(t, X) + (V' - \phi)(t, X) < 0$ by Theorem 9.11 and assumption (2), so $V_\epsilon - \phi$ cannot have a global minimum in $\overline{\mathscr{D}_T}\backslash\text{supp}(\phi)$.

We have $|(t_\epsilon, X_\epsilon) - (t, X)| < \epsilon_a$ for all $\epsilon < \epsilon_c$, since the radius of B was smaller than ϵ_a. Choose an $N \in \mathbb{N}$ such that $\epsilon_n < \epsilon_c$ for all $n \geq N$. Now we have $|(t_{\epsilon_n}, X_{\epsilon_n}) - (t, X)| < \epsilon_a$ for all $n \geq N$, and we have proved that $(t_{\epsilon_n}, X_{\epsilon_n}) \to (t, X)$.

Since V_{ϵ_n} is a viscosity supersolution of (8.3) and $(t_{\epsilon_n}, X_{\epsilon_n})$ is a minimum of $V_{\epsilon_n} - \phi$, we know that
$$F\big(t_{\epsilon_n}, X_{\epsilon_n}, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t_{\epsilon_n}, X_{\epsilon_n}, \phi)\big) + \frac{1}{\epsilon_n}\max\big\{G(D_X\phi)(t_{\epsilon_n}, X_{\epsilon_n}); 0\big\} \leq 0. \quad (10.19)$$
We see that
$$F\big(t_{\epsilon_n}, X_{\epsilon_n}, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t_{\epsilon_n}, X_{\epsilon_n}, D_X\phi)\big) \to F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, D_X\phi)\big)$$
as $n \to \infty$, by Lemma 6.3, because $(t_{\epsilon_n}, X_{\epsilon_n}) \to (t, X)$, and because U and ϕ are continuous. The term $\frac{1}{\epsilon_n}\max\big\{G\big(\phi(t_{\epsilon_n}, X_{\epsilon_n})\big); 0\big\}$ is non-negative for all $n \in \mathbb{N}$, and therefore we see by (10.19) that
$$F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, D_X\phi)\big) \leq 0. \quad (10.20)$$

The term $\frac{1}{\epsilon_n}\max\big\{G\big(\phi(t_{\epsilon_n}, X_{\epsilon_n})\big); 0\big\}$ must be bounded as $n \to \infty$, as it is non-negative, $F\big(t_{\epsilon_n}, X_{\epsilon_n}, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t_{\epsilon_n}, X_{\epsilon_n}, \phi)\big)$ is bounded and (10.19) is satisfied. We cannot have $G\big(D_X\phi(t, X)\big) > 0$, because this would imply that $\frac{1}{\epsilon_n}\max\big\{G\big(D_X\phi(t_{\epsilon_n}, X_{\epsilon_n})\big); 0\big\}$ blew up when $n \to \infty$, so we have
$$G\big(D_X\phi(t, X)\big) \leq 0. \quad (10.21)$$
Inequalities (10.20) and (10.21) imply that (10.18) is satisfied. □

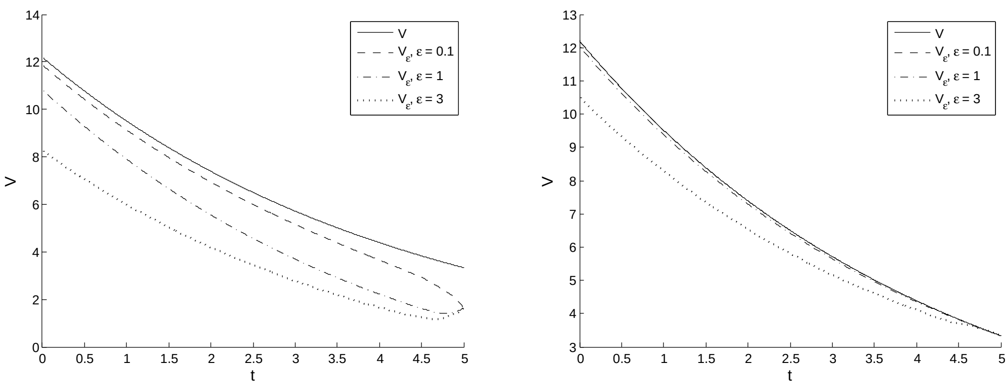
By applying the above lemma and the comparison result 6.15, we are able to prove Theorem 10.1.

Proof of Theorem 10.1: We have proved that $\{V_\epsilon\}_{\epsilon>0}$ converges uniformly to a function $V' \in C(\overline{\mathscr{D}_T})$ on $[0, T) \times \overline{\mathscr{D}}$. By Lemma 10.15, V' is a supersolution of (5.11) on $\mathscr{D}_T$, and satisfies (4.2). The value function V is a subsolution of (5.11) on $[0, T) \times \overline{\mathscr{D}}$, and satisfies (4.2). It follows by Theorem 6.15 that $V \leq V'$. On the other hand, we know that $V' \leq V$, since $V_\epsilon \leq V$ for all $\epsilon > 0$. We see that $V' = V$, i.e., V' is the unique viscosity solution of (5.11) and (4.2) in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$.

10.4 Convergence to viscosity solution, proof II

This section contains the second proof of Theorem 10.1. The difference from the proof in Section 10.3, is that we have replaced W by $\widehat{W}$ as terminal utility function, i.e., (8.5) holds instead of (8.4). This replacement makes the proof slightly easier, as we can show that $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous on $\overline{\mathscr{D}_T}$ (not only on $[0,T)\times\overline{\mathscr{D}}$), and V_ϵ converges to V on the whole domain $\overline{\mathscr{D}_T}$.

The proof becomes easier because $V(T,x,y) = V_\epsilon(T,x,y)$ for all $\epsilon > 0$ and all $(x,y)\in\overline{\mathscr{D}}$, if we replace W by $\widehat{W}$. If $W(x,y)\neq\widehat{W}(x,y)$ for some $(x,y)\in\overline{\mathscr{D}}$, and we do not replace W by $\widehat{W}$, we have $V(T,x,y)\neq V_\epsilon(T,x,y)$ and $V_\epsilon(T,x,y)\nrightarrow V(T,x,y)$. But $V_\epsilon(t,x,y)\to V(t,x,y)$ for all $t<T$, so $\{V_\epsilon\}_{\epsilon>0}$ is obviously not equicontinuous at T. If $V(T,x,y)=V_\epsilon(T,x,y)$, on the other hand, it is easy to prove equicontinuity at T, see Lemma 10.16 and Theorem 10.17 below. Figure 10.1 illustrates the difference between using W and $\widehat{W}$ as terminal utility function.



(a) W is the terminal utility function

(b) $\widehat{W}$ is the terminal utility function

Fig. 10.1: Plots of the value function V for $x=5, y=5$. The terminal time of the problem if $T=5$. When W is the terminal utility function, V_ϵ only converges to V for $t<T$. When $\widehat{W}$ is the terminal utility function, V_ϵ converges to V for all $t\in[0,T]$.

The main procedure of the proof is similar to the one in Section 10.3, and we will focus on the differences between the two proofs in this section. The first lemma we will prove is a new version of Lemma 9.17 from Section 9.2, which says that $\{V_\epsilon\}_{\epsilon>0}$ is left-equicontinuous on the whole domain $\overline{\mathscr{D}_T}$.

Lemma 10.16. *The family of functions $\{V_\epsilon\}_{\epsilon>0}$ is left-equicontinuous in t on $\overline{\mathscr{D}_T}$, i.e., for each $(t,x,y)\in\overline{\mathscr{D}_T}$, there is an $\omega:[0,T]\to\mathbb{R}$, such that $\omega(0)=0$, ω is continuous at 0 and*

$$|V_\epsilon(t,x,y)-V_\epsilon(t',x,y)|\leq\omega(|t-t'|)$$

for all $t'\in[0,t)$ and $\epsilon>0$.

Proof. Lower semi-continuity can be proved exactly as before. We also know that V_ϵ is upper semi-continuous from the left uniformly in $\epsilon>0$ for $t<T$, so what is left to

prove, is that V_ϵ is upper semi-continuous, uniformly in ϵ, for $t=T$. We want to prove that, given any $\epsilon'>0$, there is a $\varsigma>0$ such that

$$V_\epsilon(t,X)-V_\epsilon(T,X)<\epsilon'$$

for all $\epsilon>0$ and $t\in(T-\varsigma,T)$. As in the proof of Lemma 9.17, the problem can be reduced to showing that there is a $\varsigma>0$ such that

$$V_\epsilon(T,x-c,y+\beta c)-V_\epsilon(T,x,y)<\epsilon'$$

for all $c\in\big[0,\min\{x;\,\varsigma/\epsilon\}\big]$, $t\in(T-\varsigma,T)$ and $\epsilon>0$. This result obviously holds, as

$$\begin{aligned}V_\epsilon(T,x-c,y+\beta c)&=\widehat{W}(x-c,y+\beta c)\\&\le\widehat{W}(x,y)\\&=V_\epsilon(T,x,y),\end{aligned}$$

and the lemma is proved. □

Using the lemma above instead of Lemma 9.17, we can prove that V_ϵ is equicontinuous at the whole domain $\overline{\mathscr{D}_T}$. It corresponds to Theorem 9.18 in the case where we use the original function W.

Theorem 10.17. *The family of functions V_ϵ is equicontinuous at $\overline{\mathscr{D}_T}$, i.e., for each $(t,x,y)\in\overline{\mathscr{D}_T}$, there exists a function $\omega:\overline{\mathscr{D}_T}\to\mathbb{R}$, such that*

$$|V_\epsilon(t,x,y)-V_\epsilon(t',x',y')|<\omega(|t-t'|,|x-x'|,|y-y'|)$$

for all $(t',x',y')\in\overline{\mathscr{D}_T}$.

Proof. See the proof of Theorem 9.18. □

As in the previous section, we can prove that V_ϵ is pointwise convergent, but in this section we will define V' as the limit of V_ϵ on $\overline{\mathscr{D}_T}$, instead of the limit of V_ϵ on $[0,T)\times\overline{\mathscr{D}}$.

Definition 10.18. *Define $V':\overline{\mathscr{D}_T}\to\mathbb{R}$ by*

$$V'(t,X)=\lim_{\epsilon\to0}V_\epsilon(t,X).$$

As in the previous section, we know nothing about the smoothness of V' or the nature of the convergence of V_ϵ, but by using the Arzelà-Ascoli theorem and the lemma above, we obtain the following theorem.

Theorem 10.19. *The family of functions V_ϵ converges uniformly to V' on compact subsets of $\overline{\mathscr{D}_T}$ when $\epsilon\to0$, and V' is continuous.*

Proof. See the proof of Theorem 10.13. □

We can prove that V' is a viscosity supersolution of (5.11) and (4.2) just as in the previous section. Note that we do not have to extend the domain of V' to $t=T$, or to prove that $\lim_{t\to T}V'(t,x,y)=V(T,x,y)$, as V' already is defined for $t=T$, and $V'(T,x,y)=\widehat{W}(x,y)=V(T,x,y)$ for all $(x,y)\in\overline{\mathscr{D}}$. Again the reader is referred to Section 10.3 for a proof.

Lemma 10.20. *V' is a viscosity supersolution of* (5.11) *and* (4.2)*.*

Proof. See the proof of Lemma 10.24. □

As in the previous section, we use Theorem 9.11 and Lemma 10.20 to conclude that V' is the unique viscosity solution of (5.11) and (4.2) in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$, and Theorem 10.1 is proved.

10.5 Convergence to viscosity solution, proof III

This section contains our third proof of Theorem 10.1 in the general case. The proof employs a weaker definition of subsolutions and supersolutions than the one used in other parts of the thesis, and is based upon a strong comparison principle we will not prove.

We will let the terminal condition of V_ϵ be given by $\widehat{W}$ instead of W, as this implies that $V(T,x,y) = \widehat{W}(x,y) = V_\epsilon(T,x,y)$ for all $\epsilon > 0$ and all $(x,y) \in \overline{\mathscr{D}}$. We will not apply any previously proved results of this chapter or Chapter 9, except that

(I) V_ϵ is the unique viscosity solution of (8.3) and (8.5) in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$,
(II) there is a function $\tilde{V} \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$, such that
$$0 < V_\epsilon(t,X) < \tilde{V}(t,X) \tag{10.22}$$
for all $(t,X) \in \overline{\mathscr{D}_T}$ and $\epsilon > 0$, and
(III) $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous at (T,x,y) for all $(x,y) \in \overline{\mathscr{D}}$.

The claim (I) follows from Theorems 9.5 and 9.9, while (II) follows from Lemma 9.12. The only result from this chapter or Chapter 9 we need to prove (III), is that $V_\epsilon \leq V$: On one hand, we have

$$\begin{aligned}
V_\epsilon(t,x,y) \geq & \,\mathrm{E}\left[\int_t^T e^{-\delta s} U\left(Y_s^{0,0}\right)\,ds + W(X_T^{0,0}, Y_T^{0,0})\right] \\
= & \int_t^T e^{-\delta s} U\left(ye^{-\beta(s-t)}\right)\,ds + \widehat{W}\left(xe^{\hat{r}(T-t)}, ye^{-\beta(T-t)}\right) \\
\geq & \,\widehat{W}(x,y) - \omega_{\widehat{W}}\left(xe^{\hat{r}(T-t)} - x, y - ye^{-\beta(T-t)}\right) \\
= & \,V_\epsilon(T,x,y) - \omega_{\widehat{W}}\left(xe^{\hat{r}(T-t)} - x, y - ye^{-\beta(T-t)}\right)
\end{aligned} \tag{10.23}$$

for all $(t,x,y) \in \overline{\mathscr{D}_T}$, all $\epsilon > 0$ and a modulus of continuity $\omega_{\widehat{W}}$ for $\widehat{W}$, so V_ϵ is lower semi-continuous for $t = T$, uniformly in ϵ. On the other hand, we have

$$V_\epsilon(t,x,y) \leq V(t,x,y) \leq \widehat{W}(x,y) + \omega_V\big(|T-t|,0,0\big) = V_\epsilon(T,x,y) + \omega_V\big(|T-t|,0,0\big) \tag{10.24}$$

for a modulus of continuity ω_V of V, since $V_\epsilon \leq V$. It follows that V_ϵ is upper semi-continuous at $t = T$, uniformly in ϵ, and (III) is proved.

We attack the problem from another angle than we did in Sections 10.3 and 10.4, as the proofs in these sections are based upon the monotonicity of V_ϵ in ϵ and the equicontinuity of V_ϵ. In Sections 10.3 and 10.4 we first prove that V_ϵ converge uniformly, and then that the limit function is V. In this section we first prove that $V_\epsilon \to V$ pointwise, and then that the convergence is uniform.

The idea of the proof in this section is to define functions $\underline{V}$ and $\overline{V}$ by the limsup and liminf operations, respectively. We will prove that $\underline{V}$ is a weak viscosity subsolution, and that $\overline{V}$ is a weak viscosity supersolution of (5.11) and (4.2). By the comparison principle, we will have $\underline{V} \leq \overline{V}$. On the other hand we know by definition that $\overline{V} \leq \underline{V}$, so we can conclude that $\overline{V} = \underline{V}$ and that $V' := \overline{V} = \underline{V}$ is the unique viscosity solution of (5.11) and (4.2) in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$.

First we will give the definition of weak subsolutions and supersolutions. The definition is weaker in the sense that a viscosity subsolution (supersolution) always will be a weak viscosity subsolution (supersolution), while the converse result does not hold.

Definition 10.21 (Weak subsolutions and supersolutions). *Let $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$. A function $v \in USC(\overline{\mathscr{D}_T})$ $\big(v \in LSC(\overline{\mathscr{D}_T})\big)$ is a weak subsolution (weak supersolution) of* (5.11) *in $\mathscr{O}_T$ if and only if we have, for every $(t,X) \in \mathscr{O}_T$ and $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ such that (t,X) is a global maximum (minimum) relative to $\mathscr{O}_T$ of $v - \phi$,*

$$\max\big\{G(D_X\phi);\, \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t,X,\phi)\big)\big\} \geq 0 \, (\leq 0).$$

The definition of subsolutions and supersolutions above is weaker than the definition given in Definition 6.2. However, we may give a definition of viscosity solutions based on the definition above, which is equivalent to the "old" definition of viscosity solutions.

Lemma 10.22. *The function $v : C(\overline{\mathscr{D}_T})$ is a viscosity solution of* (5.11) *if and only if it is a weak viscosity subsolution of* (5.11) *on $[0,T) \times \overline{\mathscr{D}}$, and a weak viscosity supersolution of* (5.11) *on $\mathscr{D}_T$.*

Proof. This follows directly from Definition 6.2 and Definition 10.21. □

We will assume that we have a strong comparison principle, i.e., that the following theorem holds.

Theorem 10.23 (Strong Comparison Principle). *Suppose $\underline{v} \in USC(\overline{\mathscr{D}_T})$ is a weak subsolution of* (5.11) *on $[0,T) \times \overline{\mathscr{D}}$, $\overline{v} \in LSC(\overline{\mathscr{D}_T})$ is a weak supersolution of* (5.11) *on $\mathscr{D}_T$, and that $\underline{v} \leq \overline{v}$ for $t = T$. Then $\underline{v} \leq \overline{v}$ on $\overline{\mathscr{D}_T}$.*

It is relatively easy to prove a strong comparison principle in the case of Dirichlet boundary conditions, as the proof of Theorem 6.15 can be directly adapted to this case. See [35]. It is more difficult to generalize the proof to state constraint problems, as we do not know anything about the value of $\underline{v} - \overline{v} \in USC(\overline{\mathscr{D}_T})$ at the boundary of $\mathscr{D}$. Using notation from the proof of Theorem 6.15, we see that the function Φ has a maximum somewhere in $\overline{\mathscr{D}_T}$, but the maximum value could be taken at a point in Γ where $\underline{v}$ or $\overline{v}$ is discontinuous. In [40] and [41] it is explained how one can handle the state constraint boundary condition by constructing a more appropriate test function, but we will not adapt that proof to the current problem here.

The two following lemmas show that we may make some additional assumptions on the function ϕ in Definition 10.21. The lemmas may be proved in exactly the same way as Lemmas 6.5 and 6.6.

Lemma 10.24. *The function $v \in USC(\overline{\mathscr{D}_T})$ is a viscosity subsolution of* (5.11) *on $\mathscr{O}_T \subset \overline{\mathscr{D}_T}$ if and only if*

$$\max\big\{G(D_X\phi);\ \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)\big\} \geq 0 \tag{10.25}$$

for all $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t, X) \in \overline{\mathscr{D}_T}$ that satisfies these conditions:

(1) (t, X) is a global maximum of $v - \phi$ on $\mathscr{O}_T$, and there is an $\epsilon' > 0$ such that $(v - \phi)(t', X') < (v - \phi)(t, X) - \epsilon'$ for all other maxima (t', X'),
(2) $(v - \phi)(t, X) = a(t, X)$ for some given function $a : \overline{\mathscr{D}_T} \to [0, \infty)$, and
(3) $\lim_{x\to\infty, y\to\infty} \phi(t, x, y)/(x + y) \neq 0$.

Lemma 10.25. *The function $v \in LSC(\overline{\mathscr{D}_T})$ is a viscosity supersolution of* (5.11) *on $\mathscr{O}_T \subset \overline{\mathscr{D}_T}$ if and only if*

$$\max\big\{G(D_X\phi);\ \phi_t + F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big)\big\} \leq 0 \tag{10.26}$$

for all $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t, X) \in \overline{\mathscr{D}_T}$ that satisfies these conditions:

(1) (t, X) is a global minimum of $v - \phi$ on $\mathscr{O}_T$, and there is an $\epsilon' > 0$ such that $(v - \phi)(t', X') > (v - \phi)(t, X) + \epsilon'$ for all other minima (t', X'),
(2) $(v - \phi)(t, X) = -a(t, X)$ for some given function $a : \overline{\mathscr{D}_T} \to [0, \infty)$, and
(3) ϕ has compact support.

Now we define the functions $\underline{V}$ and $\overline{V}$.

Definition 10.26. *Define $\underline{U}_\epsilon, \overline{U}_\epsilon : \overline{\mathscr{D}_T} \to \mathbb{R}$, $\underline{V} \in USC(\overline{\mathscr{D}_T})$ and $\overline{V} \in LSC(\overline{\mathscr{D}_T})$ by*

$$\underline{U}_\epsilon(t, x, y) = \sup_{\hat{\epsilon}\leq\epsilon} \big\{V_{\hat{\epsilon}}(t', x', y') \ :\ |(t, x, y) - (t', x', y')| \leq \hat{\epsilon}\big\},$$

$$\overline{U}_\epsilon(t, x, y) = \inf_{\hat{\epsilon}\leq\epsilon} \big\{V_{\hat{\epsilon}}(t', x', y') \ :\ |(t, x, y) - (t', x', y')| \leq \hat{\epsilon}\big\},$$

$$\underline{V}(t, x, y) = \lim_{\epsilon\to 0} \underline{U}_\epsilon(t, x, y)$$

and

$$\overline{V}(t, x, y) = \lim_{\epsilon\to 0} \overline{U}_\epsilon(t, x, y).$$

We call the function $\underline{V}$ for the upper semi-continuous envelope of V, and we call the function $\overline{V}$ for the lower semi-continuous envelope of V. It is explained in Chapter 6 of [15] that $\underline{V}$ and $\overline{V}$ are well-defined, and that $\underline{V} \in USC(\overline{\mathscr{D}_T})$ and $\overline{V} \in LSC(\overline{\mathscr{D}_T})$. Note that $\underline{V}$ and $\overline{V}$ are not identical to the pointwise limit superior and pointwise limit inferior, respectively, of V_ϵ, as we take supremum/infimum over a neighbourhood of each point. The proofs below would work if we had defined $\underline{V}$ and $\overline{V}$ by the pointwise limit superior and pointwise limit inferior operations. In fact, they would be easier, as we would not have to use equicontinuity to prove that $\underline{V} = \overline{V}$ for $t = T$. However, it is not as easy to prove that the functions are semi-continuous if we define them by the pointwise limsup and liminf operations, and therefore we use the definitions above.

The functions $\underline{U}_\epsilon$ and $\overline{U}_\epsilon$ will be used in several proofs below, and the following lemma states some of these functions' properties.

Lemma 10.27. *The function* $\underline{U}_\epsilon$ *decreases as* ϵ *decreases, and* $\overline{U}_\epsilon$ *increases as* ϵ *decreases. We have* $\underline{U}_\epsilon \in LSC(\overline{\mathscr{D}_T})$ *and* $\overline{U}_\epsilon \in USC(\overline{\mathscr{D}_T})$ *for all* $\epsilon > 0$.

Proof. We see directly from the definition that $\underline{U}_\epsilon$ decreases as ϵ decreases, and that $\overline{U}_\epsilon$ increases as ϵ decreases. We will only prove that $\underline{U}_\epsilon \in LSC(\overline{\mathscr{D}_T})$, as we can prove that $\overline{U}_\epsilon \in USC(\overline{\mathscr{D}_T})$ by a similar argument. Assume $\underline{U}_\epsilon \notin LSC(\overline{\mathscr{D}_T})$. Then there is an $\epsilon' > 0$ and a sequence $\{(t_n, X_n)\}_{n\in\mathbb{N}}$ in $\overline{\mathscr{D}_T}$ converging to a $(t, X) \in \overline{\mathscr{D}_T}$, such that

$$\underline{U}_\epsilon(t, X) - \underline{U}_\epsilon(t_n, X_n) > \epsilon'$$

for all $n \in \mathbb{N}$. By the definition of $\underline{U}_\epsilon$, there is an $\hat{\epsilon} \leq \epsilon$ and a $(t', X') \in \overline{\mathscr{N}}\big((t, X), \hat{\epsilon}\big)$ such that

$$\underline{U}_\epsilon(t, X) < V_{\hat{\epsilon}}(t', X') + \epsilon'/2.$$

Let $\{(t'_n, X'_n)\}$ be a sequence converging to (t', X'), such that $(t'_n, X'_n) \in \overline{\mathscr{N}}\big((t_n, X_n), \hat{\epsilon}\big)$ for all $n \in \mathbb{N}$. We have

$$\underline{U}_\epsilon(t_n, X_n) \geq V_{\hat{\epsilon}}(t'_n, X'_n),$$

so

$$V_{\hat{\epsilon}}(t', X') - V_{\hat{\epsilon}}(t'_n, X'_n) > (\underline{U}_\epsilon(t, X) - \epsilon'/2) - \underline{U}_\epsilon(t_n, X_n) > \epsilon'/2.$$

This is a contradiction to the fact that $V_{\hat{\epsilon}}$ is continuous, and we see that $\underline{U}_\epsilon \in LSC(\overline{\mathscr{D}_T})$. □

By applying Lemma 10.24, we will show that $\underline{V}$ is a weak subsolution of (5.11). We proceed similarly as in the proof of Theorem 10.15: We assume $\underline{V} - \phi$ has a global maximum at (t^*, X^*), and show that a global maximum of $V_{\epsilon_n} - \phi$ converges to (t^*, X^*) for a sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ converging to 0. Using that V_{ϵ_n} is a subsolution of the penalization problem, we use convergence arguments for G and F to show that $\underline{V}$ also is a subsolution.

Lemma 10.28. *The function* $\underline{V}$ *is a weak subsolution of* (5.11) *and* (4.2) *on* $[0, T) \times \overline{\mathscr{D}}$.

Proof. We have $\underline{V}(T, x, y) = V_\epsilon(T, x, y) = \widehat{W}(x, y)$ for all $(x, y) \in \overline{\mathscr{D}}$, since $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous at (T, x, y). What is left to prove, is therefore that $\underline{V}(t, x, y)$ is a viscosity subsolution on $[0, T) \times \overline{\mathscr{D}}$.

Let $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and $(t^*, X^*) \in [0, T) \times \overline{\mathscr{D}}$ satisfy the following conditions:

(1) (t^*, X^*) is a global maximum of $V - \phi$ on $[0, T) \times \overline{\mathscr{D}}$, and there is an $\epsilon' > 0$ such that $(V - \phi)(t, X) < (V - \phi)(t^*, X^*) - 2\epsilon'$ for all other maxima $(t, X) \in [0, T) \times \overline{\mathscr{D}}$,
(2) $(V - \phi)(t^*, X^*) = 0$, and
(3) $\lim_{x\to\infty, y\to\infty} \phi(t, x, y)/(x + y) \neq 0$.

By Lemma 10.24, $\underline{V}$ is a subsolution if we can prove that (10.25) holds for ϕ. Note that the inequality in (1) remains valid if we decrease ϵ', and we will define a sufficiently small value of ϵ' below.

By definition, there are sequences $\{\epsilon_n\}_{n\in\mathbb{N}}$ converging to 0 and $\{(t_n, X_n)\}_{n\in\mathbb{N}}$ converging to (t^*, X^*), such that

$$\underline{V}(t^*, X^*) = \lim_{n\to\infty} V_{\epsilon_n}(t_n, X_n).$$

Let (t'_n, X'_n) be a maximum of $V_{\epsilon_n} - \phi$. We know that $V_{\epsilon_n} \leq \tilde{V} \in C_{\gamma'}(\overline{\mathscr{D}_T})$ by (10.22), so $\phi - V_{\epsilon_n} \to \infty$ when $x, y \to \infty$ by (3). Therefore all maxima of $V_{\epsilon_n} - \phi$ must be on some compact domain $\mathscr{O}_T \subset \overline{\mathscr{D}_T}$, where $\mathscr{O}_T$ is independent of n.

We will show by contradiction that

$$\lim_{n\to\infty}(t'_n, X'_n) = (t^*, X^*). \tag{10.27}$$

Assume $(t'_n, X'_n) \nrightarrow (t^*, X^*)$. We can assume without loss of generality that (t'_n, X'_n) converges to $(\bar{t}, \bar{X}) \neq (t^*, X^*)$ for some $(\bar{t}, \bar{X}) \in \overline{\mathscr{D}_T}$ because $\mathscr{O}_T$ is compact: If (t_n, X_n) was not convergent, we could obtained a convergent sequence by taking a subsequence. By choosing $\epsilon' > 0$ sufficiently small, we can assume

$$(\underline{V} - \phi)(\bar{t}, \bar{X}) < (\underline{V} - \phi)(t^*, X^*) - 2\epsilon' = -2\epsilon'.$$

We see that

$$\lim_{n\to\infty} \underline{U}_{\epsilon_n}(t'_n, X'_n) \leq \underline{V}(\bar{t}, \bar{X}),$$

since $\underline{V} \in USC(\overline{\mathscr{D}_T})$. It follows that

$$\lim_{n\to\infty} \underline{U}_{\epsilon_n}(t'_n, X'_n) \leq \underline{V}(\bar{t}, \bar{X}) < \phi(\bar{t}, \bar{X}) - 2\epsilon' = \lim_{n\to\infty} \phi(t'_n, X'_n) - 2\epsilon',$$

and using $V_{\epsilon_n} \leq \underline{U}_{\epsilon_n}$, we see see that

$$V_{\epsilon_n}(t'_n, X'_n) \leq \phi(t'_n, X'_n) - \epsilon' \tag{10.28}$$

for sufficiently large n. Since $V(t^*, X^*) = \phi(t^*, X^*)$, $\lim_{n\to\infty} V_{\epsilon_n}(t_n, X_n) = V(t^*, X^*)$ and $\lim_{n\to\infty} \phi(t_n, X_n) = \phi(t^*, X^*)$,

$$V_{\epsilon_n}(t^*, X^*) > \phi(t^*, X^*) - \epsilon'$$

for all sufficiently large n. This is a contradiction to (10.28), since (t'_n, X'_n) should be a global maximum of $V_{\epsilon_n} - \phi$. We can conclude that (10.27) holds.

Since V_{ϵ_n} is a subsolution, we have

$$F\big(t'_n, X'_n, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t'_n, X'_n, \phi)\big) + \frac{1}{\epsilon_n} \max\{G\big(D_X\phi(t'_n, X'_n)\big); 0\} \geq 0. \tag{10.29}$$

By Lemma 6.3 and the continuity of $D_X\phi$ and $D_X^2\phi$, we have

$$F\big(t'_n, X'_n, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t'_n, X'_n, \phi)\big) \to F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big),$$

and

$$G\big(D_X\phi(t'_n, X'_n)\big) \to G\big(D_X\phi(t, X)\big).$$

We consider two different cases, depending on whether (1) $G\big(D_X\phi(t,X)\big) \geq 0$ or (2) $G\big(D_X\phi(t,X)\big) < 0$:

(1) If $G\big(D_X\phi(t,X)\big) \geq 0$, we see immediately that (10.25) holds holds.
(2) If $G\big(D_X\phi(t,X)\big) < 0$, we see from 10.29 that $F\big(t'_n, X'_n, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t'_n, X'_n, \phi)\big) \geq 0$ for sufficiently large n, so $F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big) \geq 0$ and (10.25) holds

In both cases we see that (10.25) holds, and the lemma is proved. □

Now we will show that $\overline{V}$ is a weak supersolution by employing Lemma 10.25 and using a similar technique as above.

Lemma 10.29. *The function $\overline{V}$ is a weak supersolution of* (5.11) *and* (4.2) *on $\mathscr{D}_T$.*

Proof. $\overline{V}(T,x,y) = V_\epsilon(T,x,y) = \widehat{W}(x,y)$ for all $(x,y) \in \overline{\mathscr{D}}$, since $\{V_\epsilon\}_{\epsilon>0}$ is equicontinuous at (T,x,y). What is left to prove, is therefore that $\overline{V}(T,x,y)$ is a viscosity supersolution on $\mathscr{D}_T$.

Let $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T}) \cap C_1(\overline{\mathscr{D}_T})$ and (t^*,X^*) satisfy the following conditions:

(1) (t^*,X^*) is a global minimum of $\overline{V} - \phi$ on $\mathscr{D}_T$, and there is an $\epsilon' > 0$ such that $(\overline{V} - \phi)(t,X) < (\overline{V} - \phi)(t^*,X^*) + 2\epsilon'$ for all other minima $(t,X) \in \mathscr{D}_T$,
(2) $(\overline{V} - \phi)(t^*,X^*) = 0$, and
(3) ϕ has compact support.

By Lemma 10.25, $\overline{V}$ is a supersolution if we can prove that (10.26) holds for ϕ. Note that the inequality in (1) remains valid if we decrease ϵ', and we will define a sufficiently small value of ϵ' below.

By definition, there are sequences $\{\epsilon_n\}_{n\in\mathbb{N}}$ converging to 0 and $\{(t_n,X_n)\}_{n\in\mathbb{N}}$ converging to (t^*,X^*), such that

$$\overline{V}(t^*,X^*) = \lim_{n\to\infty} V_{\epsilon_n}(t_n,X_n).$$

Let (t_n',X_n') be a minimum of $V_{\epsilon_n} - \phi$. Since $V_{\epsilon_n} > 0$ if $x > 0$ or $y > 0$, and ϕ has compact support, all minima of $V_{\epsilon_n} - \phi$ must be on some compact domain $\mathscr{O}_T \subseteq \overline{\mathscr{D}_T}$, where $\mathscr{O}_T$ is independent of n.

We will show by contradiction that

$$\lim_{n\to\infty} (t_n',X_n') = (t^*,X^*). \tag{10.30}$$

Assume $(t_n',X_n') \nrightarrow (t^*,X^*)$. We can assume without loss of generality that (t_n',X_n') converges to $(\bar{t},\bar{X}) \neq (t^*,X^*)$ for some $(\bar{t},\bar{X}) \in \overline{\mathscr{D}_T}$: If (t_n',X_n') was not convergent, we could obtained a convergent sequence by taking a subsequence. By choosing $\epsilon' > 0$ sufficiently small, we can assume

$$(\overline{V} - \phi)(\bar{t},\bar{X}) > (\overline{V} - \phi)(t^*,X^*) + 2\epsilon' = 2\epsilon'.$$

We see that

$$\lim_{n\to\infty} \overline{V}_{\epsilon_n}(t_n',X_n') \geq \overline{V}(\bar{t},\bar{X}),$$

since $\overline{V} \in LSC(\overline{\mathscr{D}_T})$. It follows that

$$\lim_{n\to\infty} \overline{U}_{\epsilon_n}(t_n',X_n') \geq \overline{V}(\bar{t},\bar{X}) > \phi(\bar{t},\bar{X}) + 2\epsilon' = \lim_{n\to\infty} \phi(t_n',X_n') + 2\epsilon',$$

and using $V_{\epsilon_n} \geq \overline{U}_{\epsilon_n}$ we see see that

$$V_{\epsilon_n}(t_n',X_n') \geq \phi(t_n',X_n') + \epsilon'$$

for sufficiently large n. Since $V(t^*,X^*) = \phi(t^*,X^*)$, $\lim_{n\to\infty} V_{\epsilon_n}(t_n,X_n) = V(t^*,X^*)$ and $\lim_{n\to\infty} \phi(t_n,X_n) = \phi(t^*,X^*)$,

$$V_{\epsilon_n}(t_n, X_n) < \phi(t_n, X_n) + \epsilon'$$

for all sufficiently large n. This is a contradiction to (10.28), since (t'_n, X'_n) should be a global minimum of $V_{\epsilon_n} - \phi$. We can conclude that (10.30) holds.

Since V_{ϵ_n} is a supersolution, we have

$$F\big(t'_n, X'_n, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t'_n, X'_n, \phi)\big) + \frac{1}{\epsilon_n}\max\big\{G\big(D_X\phi(t'_n, X'_n)\big); 0\big\} \leq 0. \quad (10.31)$$

By Lemma 6.3 and the continuity of $D_X\phi$ and $D_X^2\phi$, we have

$$F\big(t'_n, X'_n, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t'_n, X'_n, \phi)\big) \to F\big(t, X, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t, X, \phi)\big),$$

and

$$G\big(D_X\phi(t'_n, X'_n)\big) \to G\big(D_X\phi(t, X)\big).$$

We cannot have $G\big(D_X\phi(t, X)\big) > 0$, as this would imply that the left-hand side of (10.31) converge to ∞ as $n \to \infty$, so

$$G\big(D_X\phi(t, X)\big) \leq 0. \quad (10.32)$$

Since $\frac{1}{\epsilon_n}\max\big\{G\big(D_X\phi(t'_n, X'_n)\big); 0\big\}$ is non-negative and (10.31) holds for all $n \in \mathbb{N}$, we see that

$$F\big(t'_n, X'_n, D_X\phi, D_X^2\phi, \mathscr{I}^\pi(t'_n, X'_n, \phi)\big) \leq 0. \quad (10.33)$$

It follows that (10.26) holds, and the lemma is proved. □

We will now use the strong comparison principle to prove that $V = \overline{V} = \underline{V}$, which will implies that V_ϵ converges pointwise to V.

Lemma 10.30. *The family of functions $\{V_\epsilon\}_{\epsilon>0}$ converges pointwise to V on $\overline{\mathscr{D}_T}$ as $\epsilon \to 0$.*

Proof. We know that $\underline{V}(T, x, y) = \widehat{W}(x, y) = \overline{V}(T, x, y)$ for all $(x, y) \in \overline{\mathscr{D}}$ by Lemmas 10.28 and 10.29. Lemma 10.28, Lemma 10.29 and the strong comparison principle Theorem 10.23 therefore imply that

$$\underline{V}(t, x, y) \leq \overline{V}(t, x, y)$$

for all $(t, x, y) \in \overline{\mathscr{D}_T}$. On the other hand, we see by the definition of $\underline{V}$ and $\overline{V}$ that $\underline{V} \geq \overline{V}$ on $\overline{\mathscr{D}_T}$. We can conclude that there is a function $V' : \overline{\mathscr{D}_T} \to \mathbb{R}$ that satisfies

$$V' = \underline{V} = \overline{V},$$

and we see that

$$V' = \lim_{\epsilon \to 0} V_\epsilon$$

and $V' \in C(\overline{\mathscr{D}_T})$. It follows from Lemma 10.22 that V' is a viscosity solution of (5.11) and (4.2), and $V' \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$ as $V' \leq \tilde{V} \in C'_{\gamma^*}(\overline{\mathscr{D}_T})$. By uniqueness of viscosity solutions in $C'_{\gamma^*}(\overline{\mathscr{D}_T})$ (Theorem 6.16), we can conclude that $V' = V$. □

The only result that is left to prove, is that the convergence of V_ϵ to V is uniform on compact subsets. Note that this does not follow directly from the continuity of V_ϵ and V. We will prove the result by using that V can be expressed by the limsup and liminf operations applied on V_ϵ.

Proof of Theorem 10.1: We want to prove that the convergence of V_ϵ to V is uniform. Let $\mathscr{O}_T \subset \overline{\mathscr{D}_T}$ be compact. We want to prove that, given any $\epsilon' > 0$, there is an $\epsilon^* > 0$ such that

$$(V_\epsilon - V)(t,X) < \epsilon' \quad \text{and} \quad (V - V_\epsilon)(t,X) > \epsilon'$$

for all $\epsilon \in (0,\epsilon^*)$ and $(t,X) \in \mathscr{O}_T$.

We will only prove that $(V_\epsilon - V)(t,X) < \epsilon'$, as $(V - V_\epsilon)(t,X) > \epsilon'$ can be proved by the exact same argument. The result will be proved by contradiction. Assume there is an $\epsilon' > 0$ and sequences $\{(t_n,X_n)\}_{n\in\mathbb{N}}$ and $\{\epsilon_n\}_{n\in\mathbb{N}}$ with $\epsilon_n \to 0$, such that

$$(V_{\epsilon_n} - V)(t_n,X_n) > \epsilon' \tag{10.34}$$

for all $n \in \mathbb{N}$. Since $\mathscr{O}_T$ is compact, we may assume $\{(t_n,X_n)\}$ is convergent. Define $(t^*,X^*) = \lim_{n\to\infty}(t_n,X_n)$. We may also assume that $\{\epsilon_n\}_{n\in\mathbb{N}}$ and $|(t_n,X_n) - (t^*,X^*)|$ are decreasing.

Define

$$\widehat{\underline{U}}_{\epsilon_n}(t,X) = \sup_{\epsilon\le\epsilon_n} \big\{ V_\epsilon(t',X') \,:\, |(t,X)-(t',X')| \le |(t_n,X_n)-(t^*,X^*)| \big\},$$

$$\widehat{\overline{U}}_{\epsilon_n}(t,X) = \inf_{\epsilon\le\epsilon_n} \big\{ V_\epsilon(t',X') \,:\, |(t,X)-(t',X')| \le |(t_n,X_n)-(t^*,X^*)| \big\},$$

$$\widehat{\underline{V}}(t,X) = \lim_{n\to\infty} \widehat{\underline{U}}_{\epsilon_n}(t,X)$$

and

$$\widehat{\overline{V}}(t,X) = \lim_{n\to\infty} \widehat{\overline{U}}_{\epsilon_n}(t,X).$$

We can repeat the proofs above with $\widehat{\underline{U}}_{\epsilon_n}, \widehat{\overline{U}}_{\epsilon_n}, \widehat{\underline{V}}$ and $\widehat{\overline{V}}$ instead of $\underline{U}_{\epsilon_n}, \overline{U}_{\epsilon_n}, \underline{V}$ and $\overline{V}$, and therefore we see that $\widehat{\underline{V}} = V$. By the definition of $\widehat{\underline{U}}_{\epsilon_n}$, we have

$$\widehat{\underline{U}}_{\epsilon_n}(t^*,X^*) \ge V_{\epsilon_n}(t_n,X_n)$$

for all $n \in \mathbb{N}$. For sufficiently large n, we have

$$\widehat{\underline{V}}(t^*,X^*) \ge \widehat{\underline{U}_{\epsilon_n}}(t^*,X^*) - \epsilon'/2.$$

It follows that

$$V(t^*,X^*) = \widehat{\underline{V}}(t^*,X^*) \ge \widehat{\underline{U}_{\epsilon_n}}(t^*,X^*) - \epsilon'/2 \ge V_{\epsilon_n}(t_n,X_n) - \epsilon'/2$$

for sufficiently large n. This is a contradiction to (10.34), because $V(t_n,X_n) \to V(t^*,X^*)$. We conclude that the convergence of V_ϵ to V is uniform, and the theorem is proved. $\square$

10.6 Rate of Convergence

In this section we will find an estimate for the rate of convergence of V_ϵ to V. We will not give a proof for the general case, only under the following assumptions:

(1) V is a classical solution of the terminal value problem (5.11) and (4.2), and V_ϵ is a classical solution of the terminal value problem (8.3) and (8.4) for all $\epsilon > 0$. All derivatives of V_ϵ converge to the corresponding derivatives of V when $\epsilon \to 0$.
(2) There is constant $R > 0$ and an $\epsilon^* > 0$ such that

$$\left|(V_\epsilon)_t + F\big(t, X, D_X V_\epsilon, D_X^2 V_\epsilon, \mathscr{I}^\pi(t, X, V_\epsilon)\big)\right| \le R(1 + x + y)^\gamma, \tag{10.35}$$

for all $(t, X) \in \mathscr{D}_T$ and $\epsilon < 2\epsilon^*$.

Provided these assumptions hold, we will prove that the difference between V and V_ϵ in the L_∞ norm is at most linear on compact subsets of $[0, T) \times \overline{\mathscr{D}_T}$. We will let the terminal condition of V_ϵ be given by (8.4). If we replace W by $\widehat{W}$, the proof can be done similarly, except that the rate of convergence is valid on compact subsets of $\overline{\mathscr{D}_T}$, not only on compact subsets of $[0, T) \times \overline{\mathscr{D}}$.

There are of course many cases where assumptions (1)-(2) do not hold. However, estimates holding for classical solutions of HJB equations often also hold for general solutions, and assumption (2) can be expected to hold in many cases. Since V_ϵ is of order $O\big((1 + x + y)^\gamma\big)$ (Lemma 9.3), and V_ϵ is concave and increasing (Lemma 9.1), we must have $|V_x(t, X)| \le R'(x+y)^{\gamma-1}$, $|V_y(t, X)| \le R'(x+y)^{\gamma-1}$ and $|V_{xx}(t, X)| \le R'(x+y)^{\gamma-2}$ for some constant $R' > 0$ and large X. Therefore the estimate (10.35) holds for fixed ϵ. In the proof below we need a result saying that the constant R can be chosen independently of ϵ. We know that the left-hand side of (10.35) converges uniformly to the corresponding expression with V on compact subsets of $[0, T) \times \overline{\mathscr{D}}$, but we have no results of uniformity on the whole domain $[0, T) \times \overline{\mathscr{D}}$. We note that (10.35) holds for all $\epsilon < \epsilon^*$ in the special case where the left-hand side of (10.35) is monotone (either increasing or decreasing) in ϵ.

To prove our result, we will first find an estimate for the difference between $V_\epsilon(t, X)$ for two values of ϵ. Then we will use this estimate to prove that the difference between V_ϵ and V is at most linear in ϵ. We will use the comparison principle of Theorem 9.8 to show that $V_\epsilon - V_{2\epsilon}$ is bounded by $L\epsilon$ for some $L > 0$. The idea is to prove that $V_{2\epsilon} + g_\epsilon$ is a viscosity supersolution of (8.3) with penalty parameter ϵ for some function g_ϵ and all $\epsilon > 0$, where the functions g_ϵ are bounded uniformly in ϵ on some bounded subset $\mathcal{D}_T \subset [0, T) \times \overline{\mathscr{D}}$.

Lemma 10.31. *Assume $\mathcal{D}_T \subset \overline{\mathscr{D}_T}$ is bounded. Then there are functions $w_1, w_2 \in C^\infty(\overline{\mathscr{D}_T})$ such that, for all $\epsilon < \epsilon^*$,*

(1) $V_{2\epsilon} + \epsilon w_1 + w_2$ is a viscosity supersolution of the penalty equation (8.3) *with penalty parameter ϵ, and*
(2) $w_2 \equiv 0$ in $\mathcal{D}_T$.

Proof. We define $w_1 \in C^\infty(\overline{\mathscr{D}_T})$ by

$$w_1(t, x, y) = K e^{-\delta t}(1 + \chi^{\bar{\gamma}}),$$

where χ and $\bar{\gamma}$ are defined as in Lemma 6.10, and $K > 0$ is some large constant that will be defined below. Proceeding as in the proof of Lemma 6.10, we see that w_1 is a viscosity supersolution of (5.11) for sufficiently large values of K.

We define $w_2 \in C^\infty(\mathscr{D}_T)$ and $\mathcal{D}'_T$ to be a function and a set, respectively, satisfying

(I) $\mathcal{D}_T \subset \mathcal{D}'_T \subset \mathscr{D}_T$ and $\mathcal{D}'_T$ is bounded,
(II) $w_2 \equiv 0$ in $\mathcal{D}_T$,
(III) $w_2 = Ke^{-\delta t}\big(1 + (\beta x + y)\big)^\gamma$ outside $\mathcal{D}'_T$, and
(IV) $(w_2)_t + F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) \leq e^{-\delta t}U(y)$, $G(D_X w_2) \leq 0$ everywhere.

It is easy to find w_2 and $\mathcal{D}'_T$ satisfying (I)-(III). Since the inequalities mentioned in (IV) are satisfied by $w_2 = 0$ and $w_2 = Ke^{-\delta t}\big(1 + (\beta x + y)\big)^\gamma$, we see that we can find w_2 and $\mathcal{D}'_T$ satisfying (I)-(IV).

We want to show that

$$\widehat{V}_t + F\big(t, X, D_X \widehat{V}, D_X^2 \widehat{V}, \mathscr{I}^\pi(t, X, \widehat{V})\big) + \frac{1}{\epsilon} \max\big\{G(D_X \widehat{V}); 0\big\} \leq 0, \tag{10.36}$$

where $\widehat{V} = V_{2\epsilon} + \epsilon w_1 + w_2$. First suppose $G(D_X \widehat{V}) \leq 0$. Using (IV), that $V_{2\epsilon}$ is a classical solution of (8.3) with penalty parameter 2ϵ, and that w_1 is a supersolution of (5.11), we get

$$\begin{aligned}
\widehat{V}_t + F\big(t,&X, D_X \widehat{V}, D_X^2 \widehat{V}, \mathscr{I}^\pi(t, X, \widehat{V})\big) + \frac{1}{\epsilon} \max\big\{G(D_X \widehat{V}); 0\big\} \\
&= \widehat{V}_t + F\big(t, X, D_X \widehat{V}, D_X^2 \widehat{V}, \mathscr{I}^\pi(t, X, \widehat{V})\big) \\
&\leq (V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big) \\
&\quad + \epsilon\Big((w_1)_t + F\big(t, X, D_X w_1, D_X^2 w_1, \mathscr{I}^\pi(t, X, w_1)\big)\Big) \\
&\quad + (w_2)_t + F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) - 2e^{-\delta t}U(y) \\
&= -\frac{1}{\epsilon}\max\{G(V_{2\epsilon}); 0\} + \epsilon\Big((w_1)_t + F\big(t, X, D_X w_1, D_X^2 w_1, \mathscr{I}^\pi(t, X, w_1)\big)\Big) \\
&\quad + (w_2)_t + F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) - 2e^{-\delta t}U(y) \\
&\leq 0,
\end{aligned}$$

so (10.36) holds. Now suppose $G(D_X \widehat{V}) \geq 0$. We get

$$\begin{aligned}
\widehat{V}_t + F\big(t,&X, D_X \widehat{V}, D_X^2 \widehat{V}, \mathscr{I}^\pi(t, X, \widehat{V})\big) + \frac{1}{\epsilon} \max\big\{G(D_X \widehat{V}); 0\big\} \\
&= \widehat{V}_t + F\big(t, X, D_X \widehat{V}, D_X^2 \widehat{V}, \mathscr{I}^\pi(t, X, \widehat{V})\big) + \frac{1}{\epsilon} G(D_X \widehat{V}) \\
&\leq \Big((V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big) + \frac{1}{2\epsilon} G(D_X V_{2\epsilon})\Big) \\
&\quad + \epsilon\Big((w_1)_t + F\big(t, X, D_X w_1, D_X^2 w_1, \mathscr{I}^\pi(t, X, w_1)\big)\Big) \\
&\quad + (w_2)_t + F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) - 2e^{-\delta t}U(y) \\
&\quad + \frac{1}{2\epsilon} G(D_X V_{2\epsilon}) + G(D_X w_1) + \frac{1}{\epsilon} G(D_X w_2) \\
&\leq (w_2)_t + F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) - e^{-\delta t}U(y) \\
&\quad - \Big((V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big)\Big) + G(D_X w_1).
\end{aligned}$$

We obtained the last inequality by using $G(D_X w_2) \leq 0$,

$$(w_1)_t + F\big(t, X, D_X w_1, D_X^2 w_1, \mathscr{I}^\pi(t, X, w_1)\big) \leq 0,$$

and

$$\frac{1}{2\epsilon} G(D_X V_{2\epsilon}) \leq -\Big((V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big)\Big).$$

The value of $(V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big)$ is uniformly bounded in ϵ on $\mathcal{D}'_T$ for $\epsilon < \epsilon^*$ by (2). Using this result, (IV) and equation (6.17), we see that the following inequalities are valid on $\mathcal{D}'_T$ for sufficiently large K:

$$\begin{aligned}
(w_2)_t + & F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) - e^{-\delta t} U(y) \\
& - \Big((V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big)\Big) + G(D_X w_1) \\
\leq & - \Big((V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big)\Big) - e^{-\delta t} \frac{K\bar{\gamma}}{2} \chi^{\bar{\gamma}-1} \\
\leq & \, 0.
\end{aligned}$$

By a similar technique as in the proof of Lemma 6.10, we get

$$(w_2)_t + F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) \leq e^{-\delta t}\Big(U(y) - K\delta - K\big(\delta - k(\bar{\gamma})\big)\chi^{\bar{\gamma}}\Big).$$

For $(t, X) \in \overline{\mathscr{D}_T} \backslash \mathcal{D}'_T$, we use this inequality, (III), (2) and $G(D_X w_1) \leq 0$ to get

$$\begin{aligned}
(w_2)_t + & F\big(t, X, D_X w_2, D_X^2 w_2, \mathscr{I}^\pi(t, X, w_2)\big) - e^{-\delta t} U(y) \\
& - \Big((V_{2\epsilon})_t + F\big(t, X, D_X V_{2\epsilon}, D_X^2 V_{2\epsilon}, \mathscr{I}^\pi(t, X, V_{2\epsilon})\big)\Big) + G(D_X w_1) \\
\leq & \, e^{-\delta t}\Big(U(y) - K\delta - K\big(\delta - k(\bar{\gamma})\big)\chi^{\bar{\gamma}}\Big) + C(1 + x + y)^\gamma \\
< & \, 0,
\end{aligned}$$

where the last inequality follows by choosing a sufficiently large value for K. We have proved (10.36), and the lemma follows. □

Using the lemma above, we can derive an estimate for the difference between $V_\epsilon(t, X)$ for two different values of ϵ.

Lemma 10.32. *Assume (1)-(2) hold, let $\epsilon^* > 0$, and let $\mathcal{D}_T$ be any compact subset of $\overline{\mathscr{D}_T}$. Then there is a constant $L > 0$ such that*

$$V_\epsilon(t, X) - V_{2\epsilon}(t, X) \leq L\epsilon$$

for all $\epsilon < \epsilon^*$ and $(t, X) \in \mathcal{D}_T$.

Proof. Define functions w_1 and w_2 as described in Lemma 10.31. We know that V_ϵ is a subsolution of (8.3) with penalty parameter ϵ, and that $V_{2\epsilon} + \epsilon w_1 + w_2$ is a supersolution of (8.3) with penalty parameter ϵ. We also know that

$$V_\epsilon \leq V_{2\epsilon} + \epsilon w_1 + w_2$$

for $t = T$, since $w_1, w_2 \geq 0$. It follows from the comparison principle of Theorem 9.8 that

$$V_\epsilon(t, X) \leq V_{2\epsilon}(t, X) + \epsilon w_1(t, X) + w_2(t, X)$$

for all $(t, X) \in \overline{\mathscr{D}_T}$. We define $L = \sup_{\mathcal{D}_T} w$. Since $w_2 \equiv 0$ on $\mathcal{D}_T$ for all $\epsilon > 0$, we have

$$V_\epsilon(t, X) \le V_{2\epsilon}(t, X) + L\epsilon$$

for all $(t, X) \in \mathcal{D}_T$, and the lemma follows. □

Using the lemma above, we see that the error we get when approximating V by V_ϵ, is at most linear in ϵ.

Proposition 10.33. *Assume (1)-(2) hold, and let $\mathcal{D}_T$ be any compact subset of $[0, T) \times \overline{\mathscr{D}}$. Then there is a constant $L > 0$ such that*

$$0 \le V(t, X) - V_\epsilon(t, X) < L\epsilon \tag{10.37}$$

for all $\epsilon < \epsilon^*$ and $(t, X) \in \mathcal{D}_T$.

Proof. Since V_ϵ increases when ϵ decreases, the first inequality of (10.37) follows immediately, so we only need to prove the second inequality. By Lemma 10.32, there is a constant $L' > 0$ such that

$$V_\epsilon(t, X) - V_{2\epsilon}(t, X) \le L'\epsilon$$

for all $\epsilon < \epsilon^*$ and $(t, X) \in \mathcal{D}_T$. Let $\epsilon'' > 0$. Let $N \in \mathbb{N}$ satisfy $\epsilon'/2^N < \epsilon''$. We have

$$V_{\epsilon'/2^n}(t, X) - V_{\epsilon'/2^{n-1}}(t, X) \le L'\epsilon'/2^n \tag{10.38}$$

for all $n \in \{1, \ldots, N\}$. Adding (10.38) for $n = 1, \ldots, N$, we get

$$V_{\epsilon'/2^N}(t, X) - V_{\epsilon'}(t, X) \le L' \sum_{n=1}^{N} \epsilon'/2^n < L'\epsilon'.$$

By Theorem 9.11, we know that $V_{\epsilon''}(t, X) \le V_{\epsilon'/2^N}(t, X)$, and therefore

$$V_{\epsilon''}(t, X) - V_{\epsilon'}(t, X) < L'\epsilon'. \tag{10.39}$$

Since $\epsilon'' > 0$ was arbitrary, (10.39) holds for all $\epsilon'' > 0$ and $\epsilon' < \epsilon^*$. Define $L = 2L'$, and pick any $\epsilon' < \epsilon^*$. Since $\{V_\epsilon\}_{\epsilon>0}$ converges uniformly to V on $\mathcal{D}_T$, there is an $\epsilon'' < \epsilon'$ such that

$$|V(t, X) - V_{\epsilon''}(t, X)| < L'\epsilon' \tag{10.40}$$

for all $(t, X) \in \mathcal{D}_T$. Using (10.39) and (10.40), we get

$$\begin{aligned} V(t, X) - V_{\epsilon'}(t, X) &= \big(V(t, X) - V_{\epsilon''}(t, X)\big) + \big(V_{\epsilon''}(t, X) - V_{\epsilon'}(t, X)\big) \\ &\le L'\epsilon' + L'\epsilon' \\ &< L\epsilon' \end{aligned}$$

for all $(t, X) \in \mathcal{D}_T$. Since ϵ' was arbitrary, we have proved the lemma. □

If we had assumed the terminal value of V_ϵ was given by (8.5) instead of (8.4), we would define the domain $\mathcal{D}_T$ in Proposition 10.33 as a compact subset of $\overline{\mathscr{D}_T}$ instead of a compact subset of $[0, T) \times \overline{\mathscr{D}}$, and we would manage to prove a linear error in ϵ on $\mathcal{D}_T$. We cannot let $\mathcal{D}_T$ be a general subset of $\overline{\mathscr{D}_T}$ for the terminal condition (8.4), as we need a result saying that V_ϵ converges uniformly to V on $\mathcal{D}_T$, see equation (10.40) of the proof.

Part III
Numerical approximation to the viscosity solution

In this part we are solving the penalized HJB equation (8.3) numerically. As proved in Chapter 10, the simulated solution will be a good approximation to our singular control problem for small values of ϵ.

The HJB equation of the penalty problem is discretized using a finite difference scheme. As the terminal value of V_ϵ is known, we are iterating backwards in time over the time interval $[0, T]$. The discretization of the HJB operator is from [12], while several different boundary conditions are discussed and compared. We are also deriving a penalty approximation and the corresponding finite difference scheme for the one-dimensional problem described in Section 5.1. See Chapter 11 for a description of the two-dimensional scheme, and see Chapter 13 for a description of the one-dimensional scheme. A convergence result for the two-dimensional scheme is given in Chapter 12, and our proof uses techniques and results from [5], [4] and [15].

Two challenging aspects of the implementation is the calculation of the optimal control in each time step, and the calculation of the integral term of the HJB equation. Our approach to these two challenges is discussed in Chapter 14. Chapter 14 also contains a description of various technical aspects of the implementation, in addition to a discussion of running time and realistic parameter values.

Chapter 15 contains all our simulation results. The value function and optimal controls calculated by our program are given, and we see that our simulated solution corresponds well with the explicit solution formula derived in Section 7.1. We compare the performance of different boundary conditions, discuss the impact of the various parameters and constants of the problem, compare the singular problem to the non-singular problem, study the time complexity of the schemes, and try to estimate an approximate rate of convergence.

Chapter 11
A finite difference scheme for the penalty approximation

In this section we will describe a finite difference scheme for the penalty approximation (8.3) of (5.11). We fix $\epsilon > 0$, and want to find an approximate discrete solution $\tilde{v}$ to V_ϵ.

First we will write our penalty problem on a more compact form, where the HJB equation and boundary conditions are incorporated in the same equation. The purpose is partly to simplify notation, and partly to introduce the necessary notation needed to prove consistency and convergence of our numerical scheme in Chapter 12.

Our numerical scheme is similar to the scheme described in [12], where the authors construct a finite difference scheme for a non-linear degenerate parabolic integro-PDE. However, the scheme in [12] is defined for problems on $[0, T] \times \mathbb{R}^n$, while we consider a problem defined on $\overline{\mathscr{D}_T}$, so we also need to take care of boundary conditions. It is not described in [12] how to restrict the computational domain to a bounded subset of $\mathbb{R}^n$, and we will discuss this problem here.

11.1 Alternative formulation of the penalty problem

In this section we will express our penalty problem from Chapter 8 as a single equation $\mathcal{R}(t, X, v) = 0$. This equation incorporates both the HJB equation, the terminal value condition and the boundary conditions. The approach is inspired by results in [5] and [15].

First we will write the HJB equation (8.3) on a more convenient form. We see that (8.3) can be written as

$$\sup_{a \in \mathcal{A}} \mathcal{S}^a(t, X, v) = 0, \tag{11.1}$$

where

$$\begin{aligned}
\mathcal{S}^a(t, X, v) &= v_t(t, X) + f(t, X) + \mathcal{L}^a[v](t, X) + \mathcal{I}^a[v](t, X), \\
\mathcal{L}^a[v](t, X) &= \frac{1}{2}(\sigma\pi x)^2 v_{xx} + \big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big) x v_x - \beta y v_y + c(\beta v_y - v_x), \\
\mathcal{I}^a[v](t, X) &= \mathscr{I}^a(t, X, v) \\
&= \int_{\mathbb{R}\setminus\{0\}} v(t, x + \eta^a(x, z), y) - v(t, x, y) - \eta^a(x, z) v_x(t, x, y)\, \nu(dz), \\
\eta^a(x, z) &= x\pi(e^z - 1), \qquad (11.2) \\
f(t, X) &= e^{-\delta t} U(y),
\end{aligned}$$

$$a = (\pi, c)$$

and

$$\mathcal{A} = [0, 1] \times [0, 1/\epsilon].$$

As we have seen before, (11.1) may be solved explicitly for c, but we choose to write the system on the above form, as it makes the theoretical analysis of the scheme easier. Note that we write v instead of v_ϵ for simplicity.

We will assume the terminal condition of V_ϵ is $\widehat{W}$ instead of W, i.e., the terminal condition of V_ϵ is given by (8.5) instead of (3.3). We make this assumption as we want $\tilde{v}$ to converge to V on the whole domain $\overline{\mathscr{D}_T}$ when we let $\epsilon \to 0$, see Section 10.4. At the boundary $x = 0$, the value function satisfies the following equation

$$v(t, x, y) = g(t, y), \tag{11.3}$$

where

$$g(t, y) = \int_t^T e^{-\delta s} U\left(y_j e^{-\beta(s-t)}\right) ds + W\left(0, y_j e^{-\beta(T-t)}\right),$$

see (4.1). At the boundary $y = 0$, the only boundary condition we have, is the viscosity subsolution inequality (9.1).

Using the notation introduced above, we are ready to define the function $\mathcal{R}$:

$$\mathcal{R}(t, X, v) = \begin{cases} \sup_{a \in \mathcal{A}} \mathcal{S}^a(t, X, v) & \text{if } (t, X) \in \mathscr{D}_T, \\ \widehat{W}(X) - v(t, X) & \text{if } t = T, \\ g(t, y) - v(t, X) & \text{if } x = 0, \\ p - v(t, X) & \text{if } y = 0, \end{cases} \tag{11.4}$$

where $p = \infty$. Note that the sign of $\mathcal{R}$ at the boundary of $\mathcal{D}_T$ is different for our equation and the corresponding equations in [5], [4] and [15]. The reason for this, is that we have used different signs in our definition of viscosity solutions, compared to what is done in [5], [4] and [15].

We will see that the terminal value problem (8.3) and (11.3) can be written as $\mathcal{R}(t, X, v) = 0$. First we define what we mean by a viscosity solution of $\mathcal{R}(t, X, v) = 0$.

Definition 11.1 (Viscosity Solution).

(1) A function $v \in USC(\overline{\mathscr{D}_T})$ ($v \in LSC(\overline{\mathscr{D}_T})$) is a viscosity subsolution (supersolution) of

$$\mathcal{R}(t, X, v) = 0 \tag{11.5}$$

if, for all $\phi \in C^{1,2,1}(\overline{\mathscr{D}_T})$ and $(t, X) \in \overline{\mathscr{D}_T}$ such that $v - \phi$ has a local maximum at (t, X),

$$\overline{\mathcal{R}}(t, X, v) \geq 0 \quad (\underline{\mathcal{R}}(t, X, v) \leq 0).$$

(2) A function $v \in C(\overline{\mathscr{D}_T})$ is a viscosity solution of (11.5) *if and only if it is a viscosity subsolution and a viscosity supersolution of* (11.5) *on $\overline{\mathscr{D}_T}$.*

The two functions $\underline{\mathcal{R}}$ and $\overline{\mathcal{R}}$ are defined by

$$\underline{\mathcal{R}}(t,X,v) = \liminf_{(s,Y)\to(t,X)} \mathcal{R}(s,Y,v),$$
$$\overline{\mathcal{R}}(t,X,v) = \limsup_{(s,Y)\to(t,X)} \mathcal{R}(s,Y,v).$$

We say that $\underline{\mathcal{R}}$ is the lower semi-continuous envelope of $\mathcal{R}$, and that $\overline{\mathcal{R}}$ is the upper semi-continuous envelope of $\mathcal{R}$. We see immediately that two functions $\underline{\mathcal{R}}$ and $\overline{\mathcal{R}}$ are identical to $\mathcal{R}$ in the interior of the domain $\mathscr{D}_T$. On the boundaries of the domain, $\underline{\mathcal{R}}$ and $\overline{\mathcal{R}}$ can be expressed by the min and max operators, respectively. For the boundary $x = 0$, $y > 0$ and $t < T$, for example, we have

$$\underline{\mathcal{R}}(t,X,v) = \min\left\{\sup_{a\in\mathcal{A}} \mathcal{S}^a(t,X,v);\; g(t,y) - v(t,X)\right\}$$

and

$$\overline{\mathcal{R}}(t,X,v) = \max\left\{\sup_{a\in\mathcal{A}} \mathcal{S}^a(t,X,v);\; g(t,y) - v(t,X)\right\}.$$

The definition of viscosity solutions given by Definition 11.1, is equivalent to the definition of viscosity solutions given by Definition 9.4.

Lemma 11.2. *A function v is a viscosity solution of* (8.3) *and* (8.5) *if and only if it is a viscosity solution of* (11.5).

Proof. It is shown in [4] how one can write an HJB equations with Diriclet boundary condition on the form (11.5), i.e., that one can write the complete boundary value problem as a single equation by using semi-continuous envelopes. In [15] it is explained that boundary conditions given by a subsolution inequality, can be expressed as a Dirichlet boundary condition equal to $-\infty$. See the discussion after this proof.

By the discussion in [15], we see that the boundary conditions expressed in equation (11.4), correspond to the boundary conditions given in Definition 9.4 for the boundary $y = 0$. For the boundary $x = 0$, however, we have used a Dirichlet boundary condition in our definition (11.4) of $\mathcal{R}$, while the actual boundary condition given in Definition 9.4 is a viscosity subsolution inequality. We need to check that these two boundary conditions at $x = 0$ are equivalent. We call v a *Dirichlet viscosity solution* if it satisfies the conditions of Definition 9.4, except that we have replaced the subsolution boundary condition at $x = 0$ by (11.3). We need to prove that this definition is equivalent to Definition 9.4. We know that the unique viscosity solution of (8.3) and (8.5) satisfies the boundary condition (11.3), so we have proved that a viscosity solution v also is a Dirichlet viscosity solution. We can prove uniqueness of Dirichlet viscosity solutions by techniques in the proof of Theorem 9.8. By existence and uniqueness results for viscosity solutions, we see that a function v is a viscosity solution of (8.3) and (8.5) if and only if it is a Dirichlet viscosity solution. □

In the proof above, we mentioned that a subsolution boundary condition can be expressed as a Dirichlet boundary condition equal to $-\infty$, when we construct the function $\mathcal{R}$. We will explain shortly why this is the case, i.e., why the boundary conditions stated in Definition 11.1 are equivalent to the subsolution inequality (9.1) for $y = 0$. We assume v is a classical solution of (11.5), but the argument can easily be generalized to a viscosity solution v. If v is a classical solution of (11.5), we have

$$\underline{\mathcal{R}}(t, X, v) \leq 0 \tag{11.6}$$

and

$$\overline{\mathcal{R}}(t, X, v) \geq 0. \tag{11.7}$$

For $y = 0, x > 0, t < T$, (11.7) is equivalent to

$$\max\left\{\sup_{a\in\mathcal{A}} \mathcal{S}^a(t, X, v);\, p - v(t, X)\right\} \geq 0.$$

Since $p = -\infty$, this inequality implies that $\sup_{a\in\mathcal{A}} \mathcal{S}^a(t, X, v) \geq 0$, and we see that v is a subsolution at the boundary $y = 0$. Conversely, suppose v satisfies the viscosity subsolution inequality at the boundary $y = 0$. Then (11.7) is obviously satisfied. We also see that (11.6) is satisfied, as

$$\underline{\mathcal{R}}(t, X, v) = \min\left\{\sup_{a\in\mathcal{A}} \mathcal{S}^a(t, X, v);\, p - v(t, X)\right\} = -\infty.$$

11.2 Discretization of HJB equation

In this section we will describe how we discretize the HJB equation (11.5) for interior points of our computational domain. We want to find V_ϵ on the domain $\overline{\mathscr{D}_T}$, but we must restrict ourselves to a bounded domain in order to obtain a system of equations of finite dimension. We restrict ourselves to the domain

$$\mathcal{D}_T = [0, T) \times \mathcal{D} \subset \overline{\mathscr{D}_T},$$

where $\mathcal{D} = (0, x_{max}) \times (0, y_{max})$, $x_{max}, y_{max} > 0$. Let $N_t, N_x, N_y \in \mathbb{N}$, and define

$$\Delta t = \frac{T}{N_t},\ \Delta x = \frac{x_{max}}{N_x},\ \Delta y = \frac{y_{max}}{N_y}$$

and

$$t_m = m\Delta t,\ x_i = i\Delta x,\ y_j = j\Delta y,\ X_\alpha = X_{i,j} = (x_i, y_j)$$

for $m = 0, \ldots, N_t$, $i = 0, \ldots, N_x$, $j = 0, \ldots, N_y$, $\alpha = (i, j)$. We also define the following sets of grid points:

$$\begin{aligned}
\mathcal{G} &= \{(i, j) \in \mathbb{N}^2 \,:\, 0 < i < N_x,\, 0 < j < N_y\},\\
\mathcal{G}^* &= \{(i, j) \in \mathbb{N}^2 \,:\, 0 < i < N_x,\, 0 \leq j < N_y\},\\
\overline{\mathcal{G}} &= \{(i, j) \in \mathbb{Z}^2 \,:\, 0 \leq i \leq N_x,\, 0 \leq j \leq N_y\},\\
\mathcal{G}_T &= \{m \in \mathbb{N} \,:\, 0 \leq m < N_t\} \times \mathcal{G},\\
\mathcal{G}_T^* &= \{m \in \mathbb{N} \,:\, 0 \leq m < N_t\} \times \mathcal{G}^*,\\
\overline{\mathcal{G}_T} &= \{(m, i, j) \in \mathbb{Z}^3 \,:\, 0 \leq m \leq N_t,\, 0 \leq i \leq N_x,\, 0 \leq j \leq N_y\},
\end{aligned}$$

$$\mathcal{H}_T = \{(t,x,y) \in \overline{\mathcal{D}_T} \,:\, t = m\Delta t,\, x = i\Delta x,\, y = j\Delta y,\, (m,i,j) \in \mathcal{G}_T\},$$
$$\mathcal{H}_T^* = \{(t,x,y) \in \overline{\mathcal{D}_T} \,:\, t = m\Delta t,\, x = i\Delta x,\, y = j\Delta y,\, (m,i,j) \in \mathcal{G}_T^*\},$$
$$\overline{\mathcal{H}_T} = \{(t,x,y) \in \overline{\mathcal{D}_T} \,:\, t = m\Delta t,\, x = i\Delta x,\, y = j\Delta y,\, (m,i,j) \in \overline{\mathcal{G}_T}\}.$$

Each of these sets are uniquely determined by x_{max}, y_{max}, Δx, Δy and Δt. We define

$$|\mathcal{H}_T| = \max\{\Delta x, \Delta y, \Delta t\}.$$

Let $\tilde{v} : \overline{\mathcal{H}_T} \to \mathbb{R}$ denote a discrete approximation to v, and define

$$\tilde{v}_{i,j}^m = \tilde{v}(t_m, x_i, y_j) \approx v(t_m, x_i, y_j)$$

for all $(m,i,j) \in \overline{\mathcal{H}_T}$.

We discretize (11.1) by

$$\sup_{a \in \mathcal{A}} \mathcal{S}^a_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = 0, \tag{11.8}$$

where

$$\begin{aligned}\mathcal{S}^a_{\mathcal{H}}(t_m, X_\alpha, \tilde{v}) = &\frac{1}{\Delta t}\left(\tilde{v}_{i,j}^m - \tilde{v}_{i,j}^{m-1}\right) + f_{i,j}^{m-1} + \theta \mathcal{L}^a_{\mathcal{H}}[\tilde{v}]_{i,j}^{m-1} + (1-\theta)\mathcal{L}^a_{\mathcal{H}}[\tilde{v}]_{i,j}^m \\ &+ \zeta \mathcal{I}^a_{\mathcal{H}}[\tilde{v}]_{i,j}^{m-1} + (1-\zeta)\mathcal{I}^a_{\mathcal{H}}[\tilde{v}]_{i,j}^m\end{aligned}$$

for some $\theta \in [0,1]$ and $\zeta \in [0,1]$,

$$f_{i,j}^m = f(t_m, x_i, y_j) = e^{-\delta t_m} U(y_j),$$

$$\begin{aligned}\mathcal{L}^a_{\mathcal{H}}[\tilde{v}] = &\frac{1}{2}(\sigma\pi x)^2 \Delta_{xx,\Delta x}\tilde{v} + \big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big)x\delta^+_{x,\Delta x}\tilde{v}, -\beta y \delta^-_{y,\Delta y}\tilde{v} \\ &+ \frac{1}{\epsilon}\max\left\{\beta\delta^+_{y,\Delta y}\tilde{v} - \delta^-_{x,\Delta x}\tilde{v};\, 0\right\},\end{aligned}$$

and the finite differences used in the formulas above are given by

$$\delta^{\pm}_{r,h}\phi(r,t,X) = \pm\frac{1}{h}\Big(\phi(r \pm h, t, X) - \phi(r,t,X)\Big)$$

and

$$\Delta_{rr,h}\phi(r,t,X) = \frac{1}{h^2}\Big(\phi(r+h,t,X) - 2\phi(r,t,X) + \phi(r-h,t,X)\Big)$$

for functions $\phi : \mathbb{R} \times \overline{\mathscr{D}_T}$. Note that the argument r in the finite difference formulas may be equal to one of the other input variables, but also may be another variable. We will define the discretization $\mathcal{I}^a_{\mathcal{H}}$ of $\mathcal{I}^a$ in Section 11.2.1 below,

If $\theta, \zeta = 0$, the scheme is purely explicit, while $\theta, \zeta = 1$ correspond to a purely implicit scheme. Only the purely explicit scheme will be implemented, but a theoretical analysis will be performed for both schemes.

11.2.1 Discretization of the integral operator

We will use the same discretization of the integral operator $\mathcal{I}^a$ as in [12]. It is shown in [12] that

$$\mathcal{I}^a[\phi](t,X) = \mathcal{I}^{a,+}[\phi](t,X) + \mathcal{I}^{a,-}(t,X) - \tilde{b}^a(t,X)\phi_x(t,X),$$

where $\tilde{b}^a(t,X) = \int_{-\infty}^{\infty} \partial_z^2 \eta^a(x,z)\tilde{k}(z)\,dz$,

$$\mathcal{I}^{a,\pm}[\phi] = \pm \int_0^{\pm\infty} \partial_z^2[\phi(t, x+\eta^a(x,z), y)]\tilde{k}(z)\,dz$$

and

$$\tilde{k}(z) = \begin{cases} \int_{-\infty}^{z}\int_{-\infty}^{w} k(r)\,dr\,dw \text{ if } z<0, \\ \int_{z}^{\infty}\int_{w}^{\infty} k(r)\,dr\,dw \text{ if } z>0. \end{cases} \tag{11.9}$$

The discretization of $\mathcal{I}^{a,\pm}$ used in [12], is defined by

$$\mathcal{I}_{\mathcal{H}}^{a,\pm}[\phi](t,X) = \sum_{n=0}^{\infty} \Delta_{zz,\Delta z}[l_{\mathcal{H}}\phi(t, x+\eta^\alpha(x,z_n), y)]\tilde{k}_{\mathcal{H},n}^{\pm}, \tag{11.10}$$

where $z_n = n\Delta x$ (not $n\Delta z$), $\Delta z = \sqrt{\Delta x}$, $l_{\mathcal{H}}$ is a second order interpolation operator, and

$$\tilde{k}_{\mathcal{H},n}^{\pm} = \pm \int_{z_{\pm n}}^{z_{\pm(n+1)}} \tilde{k}(z)\,dz.$$

That $l_{\mathcal{H}}$ is a second order interpolation operator, means that it satisfies

$$l_{\mathcal{H}}\phi(t,x,y) = \sum_{i=0}^{N_x} w_i(x)\phi(t,x_i,y)$$

and

$$|\phi(t,x,y) - l_{\mathcal{H}}\phi(t,x,y)| \le K_I \Delta x^2 |D^2\phi|_{L^\infty}$$

for all $x \in [0, x_{max}]$ and $\phi \in C^2(\overline{\mathscr{D}_T})$, where K_I is some constant and $w_i : \mathbb{R} \to [0,1]$, $i = 0,\dots,N_x$, are basis functions satisfying $w_i(x_i) = 1$, $w_i(x_j) = 0$ for $i \neq j$ and $\sum_{i=0}^{N_x} w_i \equiv 1$. The precise definition of w_i that will be used in our implementation is:

$$w_i(x) = \begin{cases} 0 & \text{if } |x - x_i| \ge \Delta x, \\ (x - x_{i-1})/\Delta x & \text{if } x_{i-1} < x \le x_i, \\ (x_{i+1} - x)/\Delta x & \text{if } x_i < x < x_{i+1}. \end{cases}$$

We see easily that w_i satisfies the wanted properties.

The term $\tilde{b}^a(t,X)\phi_x$ is discretized by an upwind difference

$$\tilde{b}^a(t,X)\delta_{x,\Delta x}^{-}\tilde{v}.$$

Note that $\tilde{b}^a(t,X)$ is an integral that does not involve ϕ, and therefore it can calculated explicitly. If the integral cannot be solved analytically, it can be solved numerically, for example by a Newton-Cotes formula, see [19].

11.3 Boundary conditions and terminal condition

To get a complete numerical scheme, we need to specify how to find $\tilde{v}$ for the terminal time T, and at the boundaries $x = 0$, $y = 0$, $x = x_{max}$ and $y = y_{max}$. At the boundaries $x = 0$ and $y = 0$ the boundary condition of the continuous problem is given by the subsolution inequality (9.1). At the boundaries $x = x_{max}$ and $y = y_{max}$, however, the continuous problem do not have any boundary conditions, as the boundaries are artificially made in order to get a bounded domain.

We will see that it is relatively easy to find appropriate boundary conditions for $\tilde{v}$ for $t = T$ and $x = 0$. At the boundaries $y = 0$, $x = x_{max}$ and $y = y_{max}$, however, it is more difficult to find appropriate boundary conditions. We will consider several alternative boundary conditions for these boundaries, and will compare their performance in Chapter 15.

By (8.5), we define the terminal condition at $t = T$ by

$$\tilde{v}_{i,j}^{N_T} = \widehat{W}(x_i, y_j).$$

It is not obvious how to transform the subsolution boundary condition at $x = 0$ and $y = 0$ into an explicit discrete boundary condition equation. In most other articles where the solution to state constraint problems are simulated, the authors find some explicit solution formula at the boundary of the domain by studying the original optimization problem. See for example [10], [32] and [1].

We will also use this technique at the boundary $x = 0$. By (11.3), it is natural to define

$$\tilde{v}_{0,j}^{m} = g(t_m, y_j). \tag{11.11}$$

The integral in the definition of g, can be evaluated by either analytical or numerical integration.

11.3.1 Boundary condition for $y = 0$

At the boundary $y = 0$ we do not have an explicit solution formula, as we have for $x = 0$. We will consider three different strategies for determining the boundary condition at the boundary $y = 0$.

The two first strategies are based on the HJB equation of the penalty approximation. We know from Theorem 9.5 that V_ϵ satisfies the subsolution inequality (9.1) at $y = 0$. If v satisfies $v_t + F\big(t, X, D_X v, D_X^2 v, \mathscr{I}^\pi(t, X, v)\big) = 0$ at the boundary $y = 0$, we see that (8.3) is satisfied, as the term $\max\{G(D_X v;\, 0)\}$ always is positive. We obtain the following scheme:

$$\sup_{\pi \in [0,1]} \mathcal{S}_{\mathcal{H}_T}^{(\pi,0)}\big(t_m, (x_i, 0), \tilde{v}\big) = 0. \tag{11.12}$$

Note that no derivatives of y occur in this equation, as the term $-\beta y \delta_{y,\Delta y}^{-} \tilde{v}$ in $\mathcal{L}_{\mathcal{H}}^{a}$ disappears for $y = 0$. We will therefore not need any values of $\tilde{v}$ outside the grid, as we might have needed if derivatives with respect to y had occurred.

We obtain the second scheme in a similar way. The subsolution inequality (9.1) is satisfied if (8.3) holds. This suggests that we may extend the scheme (11.8) to also be valid for $j = 0$. Inserting $y = 0$, $j = 0$ into (11.8), we obtain the following scheme for determining $\tilde{v}$ at the boundary $y = 0$:

$$\sup_{a\in\mathcal{A}} \mathcal{S}^a_{\mathcal{H}_T}\big(t_m, (x_i, 0), \tilde{v}\big) = 0. \tag{11.13}$$

Again we see that no values of $\tilde{v}$ outside the grid are needed. The term $-\beta y \delta^-_{y,\Delta y}\tilde{v}$ disappears since $y = 0$, and the term βv_y is approximated by $\delta^+_{y,\Delta y}\tilde{v}$, so only values inside the grid are needed.

For other control problems it might not be possible to extend the finite difference scheme to also be valid at the boundaries of the domain. The terms involving points outside the grid, may not disappear at the boundaries. It is not possible to replace upwind differences by downwind differences, as this would give us schemes that were not monotone or stable. Therefore we consider a third alternative boundary condition.

The third alternative boundary condition is based on the HJB equation of the unpenalized problem. We know that the penalty approximation converges uniformly to the unpenalized problem as $\epsilon \to 0$, so for small ϵ it is reasonable to assume that the boundary condition for the penalized problem is approximately equal to the boundary condition for the unpenalized problem. The boundary condition at $y = 0$ for the unpenalized problem, is given by the subsolution inequality (6.2). Therefore we need to have either

$$v_t + F\big(t, X, D_X v, D_X^2 v, \mathscr{I}^\pi(t, X, v)\big) \geq 0,$$

or $G(D_X v) \geq 0$. We see that v satisfies the subsolution inequality if $G(D_X v) = 0$, i.e., if

$$\beta v_y - v_x = 0.$$

There are two reasons why we assume $G(D_X v) = 0$, instead of assuming

$$v_t + F\big(t, X, D_X v, D_X^2 v, \mathscr{I}^\pi(t, X, v)\big) = 0 \tag{11.14}$$

holds:

(1) We know that (11.14) is associated with low or zero consumption rate, while $G(D_X v) = 0$ is associated with large consumption rate or consumption gulps, see Chapter 5 and Section 7.1. It is natural to assume that the system will try to maintain a balance between the sizes of $X^{\pi,C}$ and $Y^{\pi,C}$, and therefore have a high consumption rate when $Y^{\pi,C}$ is small. Therefore we assume $G(D_X v) = 0$ for $y = 0$.
(2) In Chapter 7.1 we found an explicit solution formula for V for one specific choice of utility functions U and W. We saw that (11.14) holds when $x \leq ky$, and that $G(D_X v) = 0$ holds when $x \geq ky$. Assuming V will have a similar form for other utility functions, it is natural to assume $G(D_X v) = 0$ for $y = 0$.

When using the second method, the scheme for $y = 0$ becomes:

$$\beta \delta^+_{y,\Delta y}\tilde{v} - \delta^-_{x,\Delta x}\tilde{v} = 0. \tag{11.15}$$

We see that $\tilde{v}_{i,0}$ can be determined from $\tilde{v}_{i-1,0}$ and $\tilde{v}_{i,1}$, and that $\tilde{v}_{0,0}$ can be determined by (11.11). It follows by induction that $\tilde{v}_{i,0}$ can be determined for all $i \in \{0, \ldots, N_x\}$.

11.3.2 Boundary condition for large x and y

As mentioned above, our continuous problem does not have any boundary conditions at $x = x_{max}$ and $y = y_{max}$, as these boundaries are interior points of the domain $\mathscr{D}_T$. The ideal boundary condition would be to extend the finite difference scheme (11.8) to also be valid for $x = x_{max}$ and $y = y_{max}$, but this is not possible, as we would need grid points outside our grid. It is stated in [25] and [4] that which boundary conditions we choose at $x = x_{max}$ and $y = y_{max}$, not have any influence on the simulated solution in a fixed point $(t, X) \in \overline{\mathscr{D}_T}$ in the limiting case $x_{max}, y_{max} \to \infty$ and $|\mathcal{H}_T| \to 0$. However, a good guess on the boundary conditions at x_{max} and y_{max}, will make the convergence $\tilde{v} \to V_\epsilon$ faster.

We will consider these two boundary conditions for the boundary $x = x_{max}$:

(1) If $U_y, W_x, W_y \to 0$ as $x, y \to \infty$, it is reasonable to assume that also $V_x \to 0$ as $x \to \infty$. We obtain the Neumann boundary condition:
$$\tilde{v}^m_{N_x,j} = \tilde{v}^m_{N_x-1,j} \tag{11.16}$$
for all $m \in \{0, ..., N_t - 1\}, j \in \{1, ..., N_y\}$. The Neumann boundary condition has the advantages that it is easy to implement, and that it is relatively easy to prove stability results.

(2) As mentioned earlier, it is reasonable to assume there is consumption if x is large compared to y, and we have $G(D_X V) = 0$ if there is consumption. Assuming $G(D_X V) = 0$ for $x = x_{max}$ and $G(D_X V) \approx G(D_X V_\epsilon)$, we obtain the following scheme:
$$\beta\delta^+_{y,\Delta y}\tilde{v} - \delta^-_{x,\Delta x}\tilde{v} = 0. \tag{11.17}$$
The equation can be solved for $\tilde{v}_{N_x,j}$ in terms of $\tilde{v}_{N_x-1,j}$ and $\tilde{v}_{N_x,j+1}$. If we know the value of $\tilde{v}_{N_x,N_y}$ and $\tilde{v}_{N_x-1,j}$ for all $j \in \{0, ..., N_y\}$, we can determine $\tilde{v}_{N_x,j}$ for all $j \in \{0, ..., N_y\}$ by induction.

For $y = y_{max}$ these boundary conditions are possible:

(1) As for the boundary $x = x_{max}$, the Neumann boundary condition is possible. We define
$$\tilde{v}^{m-1}_{i,N_y} = \tilde{v}^{m-1}_{i,N_y-1} \tag{11.18}$$
for $i \in \{1, ..., N_x\}$.

(2) If y is very large compared to x, it is reasonable to assume it is not optimal to consume, see earlier discussions. If it is not optimal to consume, V_ϵ satisfies (11.14). Assuming the left-hand side of (11.14) is approximately equal when evaluated with V and V_ϵ, we see that V_ϵ approximately satisfies (11.14). Discretizing (11.14), we obtain the following scheme for the boundary $y = y_{max}$:
$$\sup_{\pi\in[0,1]} \mathcal{S}^{(\pi,0)}_{\mathcal{H}_T}\big(t_m, (x_i, y_{N_y}), \tilde{v}\big) = 0. \tag{11.19}$$

(3) If y is very large compared to x, we may assume x is negligible compared to y, and that the change in $Y^{\pi,C}$ due to consumption is negligible. One possibility is to assume $x = 0$ and use (11.11) as boundary condition for $y = y_{max}$. We will use

the somewhat better approximation $X_s^{\pi,C} = xe^{\hat{r}(s-t)}$ and $Y_s^{\pi,C} = ye^{-\beta(s-t)}$, and use a formula similar to (11.11):

$$\tilde{v}_{i,N_y}^m = \int_{t_m}^T e^{-\delta s} U\left(y_j e^{-\beta(s-t_m)}\right) ds + W\left(xe^{\hat{r}(t_m-t)}, ye^{-\beta(T-t_m)}\right). \tag{11.20}$$

11.4 The complete numerical scheme

We will employ boundary conditions (11.13), (11.17) and (11.19) in most of our simulations later. The boundary conditions (11.13), (11.17) and (11.19) are not entirely sufficient to determine $\tilde{v}$ for all boundary points. We need to determine $\tilde{v}_{N_x,N_y}$ by a separate algorithm, as neither (11.17) nor (11.19) can be used to define $\tilde{v}_{N_x,N_y}$. We define

$$\tilde{v}_{N_x,N_y}^m = \tilde{v}_{N_x-1,N_y-1}^m$$

for $m = 0, \ldots, N_t-1$, assuming V_x and V_y are small for $X = (x_{max}, y_{max})$. The complete scheme can be summarized as

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = 0, \tag{11.21}$$

for all $(m, \alpha) \in \overline{\mathcal{G}_T}$, where

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = \widehat{W}(x_i, y_j) - \tilde{v}_{i,j}^{N_t} \quad \text{for } m = N_t, \tag{11.22a}$$

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = \sup_{a \in \mathcal{A}} \mathcal{S}_{\mathcal{H}_T}^a(t_m, X_\alpha, \tilde{v}) \quad \text{for all } (m, \alpha) \in \mathcal{G}_T^*, \tag{11.22b}$$

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = g_j^m - \tilde{v}_{0,j}^m \text{ for } i = 0, \tag{11.22c}$$

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = \tilde{v}_{N_x-1,N_y-1}^m - \tilde{v}_{N_x,N_y}^m \text{ for } \alpha = (N_x, N_y), \tag{11.22d}$$

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = \sup_{\pi \in [0,1]} \mathcal{S}_{\mathcal{H}}^{\pi,0}(t_m, X_{i,N_y}, \tilde{v}) \quad \text{for } j = N_y, 0 < i < N_x, m < N_t, \tag{11.22e}$$

$$\mathcal{R}_{\mathcal{H}_T}(t_m, X_\alpha, \tilde{v}) = \beta\delta_{y,\Delta y}^+ \tilde{v}_{N_x,j}^m - \delta_{x,\Delta x}^- \tilde{v}_{N_x,j}^m = 0 \quad \text{for } i = N_x, 0 < j < N_y, m < N_t. \tag{11.22f}$$

If $\theta, \zeta = 0$, i.e., the scheme is purely explicit, $\tilde{v}_{i,j}^m$ can be determined for all $(m, i, j) \in \overline{\mathcal{G}_T}$ by the following algorithm:

(1) Determine $\tilde{v}_{i,j}^{N_t}$ for $(i, j) \in \overline{\mathcal{G}}$ by (11.22a). Set $m = N_t - 1$.
(2) Determine $\tilde{v}_{i,j}^m$ for $(i, j) \in \mathcal{G}^*$ by (11.22b).
(3) Determine $\tilde{v}_{0,j}^m$ for $j = 0, 1, \ldots, N_y$ by (11.22c).
(4) Determine $\tilde{v}_{N_x,N_y}^m$ by (11.22d).
(5) Determine $\tilde{v}_{i,N_y}^m$ for $i = 1, 2, \ldots, N_x - 1$ by (11.22e).
(6) Determine $\tilde{v}_{N_x,j}^m$ for $j = 0, 1, \ldots, N_y - 1$ by (11.22f).
(7) Stop if $m = 0$. Otherwise reduce m by 1 and go to point 2.

The order of (3)-(6) may be changed, as long as (6) comes after (5). The most challenging parts of these steps are (1) and (4). We need to determine the optimal value of the controls at each time step. This is relatively easy if $\nu \equiv 0$, but is harder for the general case $\nu \not\equiv 0$, as $\mathcal{S}$ is non-linear in π. We will discuss in Chapter 14 how to determine the optimal value of π. The idea is to assume that the optimal value of π is continuous in time, and therefore always search for a value of π close to the value of π for the previous time step.

If $\theta \neq 0$ or $\zeta \neq 0$, i.e., the scheme is not purely explicit, we need to solve a system of equations at each time step, and the scheme can be summarized as follows:

1. Determine $\tilde{v}_{i,j}^{N_t}$ for $(i,j) \in \overline{\mathcal{G}}$ by (11.22b). Set $m = N_t - 1$.
2. Determine $\tilde{v}_{i,j}^{m}$ for $(i,j) \in \overline{\mathcal{G}}$ by solving the system of equations (11.21).
3. Stop if $m = 0$. Otherwise reduce m by 1 and go to point 2.

Note that the system of equations we solve in point 2 is non-linear due the control a. The Banach fixed point theorem of the next chapter describes how we can deal with the implicitness of the scheme by an iterative method. The problem of finding an optimal control (π, C) at each time step, can be solved as for the explicit scheme. The implicit scheme will not be implemented in this thesis.

Chapter 12
Convergence of the finite difference scheme

In this chapter we will prove that the numerical scheme (11.21) is solvable, and that its solution converges to the viscosity solution of the penalty approximation, provided the grid parameters satisfy certain conditions. First we will prove that the scheme is stable, monotone and consistent, which imply that the solution of the scheme converges to the viscosity solution of the penalty approximation. Then we will use the monotonicity of the scheme and the Banach fixed point theorem to prove that the scheme has a unique, well-defined solution.

Our numerical scheme (11.21) is similar to a scheme described in [12], except that the domain in our problem is different, and that the scheme and HJB equation described in [12] are written on a more general form. Well-definiteness and the rate of convergence of the scheme is proved in [12]. We will prove that our scheme is well-defined by a similar proof as in [12]. Our proof of convergence, however, will be different, and we will use results in [5], [4] and [15], instead of using techniques from [12]. The proof of convergence in [12] uses another definition of consistency, because the domain of the value function is different, and because the objective in [12] is not only to prove convergence, but also to find a bound on the rate of convergence.

First we define what we mean by consistency, monotonicity and stability for our terminal value problem. The definitions are consistent with the definitions given in [5] and [4].

Definition 12.1 (Consistency). *The numerical scheme* (11.21) *is consistent if, for each smooth function* $\phi : \overline{\mathscr{D}_T} \to \mathbb{R}$ *and each* $(t, X) \in \overline{\mathcal{D}_T}$ *with* $x < x_{max}$ *and* $y < y_{max}$*,*

$$\liminf_{\substack{|\mathcal{H}_T| \to 0, \\ (t_n, X_\alpha) \to (t, X), \\ \xi \to 0}} \mathcal{R}_{\mathcal{H}_T}(t_n, X_\alpha, \phi + \xi) \geq \underline{\mathcal{R}}(t, X, \phi), \tag{12.1}$$

and

$$\limsup_{\substack{|\mathcal{H}_T| \to 0, \\ (t_n, X_\alpha) \to (t, X), \\ \xi \to 0}} \mathcal{R}_{\mathcal{H}_T}(t_n, X_\alpha, \phi + \xi) \leq \overline{\mathcal{R}}(t, X, \phi). \tag{12.2}$$

Note that we do not include points at the boundaries $x = x_{max}$ and $y = y_{max}$ in the definition of consistency. The boundary conditions we imposed at these boundaries were only introduced as we need to truncate the domain $\mathscr{D}_T$. As mentioned earlier, the boundary conditions at $x = x_{max}$ and $y = y_{max}$ have no effect on the simulated solution in a fixed point $(t, X) \in \overline{\mathscr{D}_T}$ in the limiting case $x_{max}, y_{max} \to \infty$ and $|\mathcal{H}_T| \to 0$. Therefore our numerical solution will converge independently of the chosen boundary condition at $x = x_{max}$ and $y = y_{max}$, and we do not have to check consistency at these boundaries.

Also note that the definition of consistency in the interior of the domain is equivalent to

$$\lim_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n, X_\alpha)\to(t,X)}} \left(\sup_{a\in\mathcal{A}} \mathcal{S}^a_{\mathcal{H}_T}(t_n, X_\alpha, \phi) - \sup_{a\in\mathcal{A}} \mathcal{S}^a(t, X, \phi) \right) = 0.$$

For problems defined on $\mathbb{R}^n$, or for problems that are not degenerate, we may define consistency by this equation instead of the more complicated definition given above, see [4].

Definition 12.2 (Monotonicity). *The numerical scheme* (11.21) *is monotone if there are positive functions $b^{k,m}_{\tilde{\alpha},\alpha} : \mathcal{A} \to [0, \infty)$ for all $\tilde{\alpha}, \alpha \in \overline{\mathcal{G}}$ and $k, m \in \{0, 1, \dots, N_t\}$, such that the scheme can be written as*

$$\sup_{a\in\mathcal{A}} \left\{ - b^{n,n-1}_{\tilde{\alpha},\tilde{\alpha}}(a)\tilde{v}^{n-1}_{\tilde{\alpha}} + \sum_{\alpha\in\overline{\mathcal{G}}\setminus\{\tilde{\alpha}\}} b^{n,n-1}_{\tilde{\alpha},\alpha}(a)\tilde{v}^{n-1}_{\alpha} + \sum_{\alpha\in\overline{\mathcal{G}}} b^{n,n}_{\tilde{\alpha},\alpha}(a)\tilde{v}^{n}_{\alpha} + \Delta t f^{a,n}_{\tilde{\alpha}} \right\} = 0 \quad (12.3)$$

for all $(n, \tilde{\alpha}) \in \overline{\mathcal{G}_T}$ and all $\tilde{v} \in \mathbb{R}^{\overline{\mathcal{H}_T}}$.

Definition 12.3 (Stability). *The numerical scheme* (11.21) *is stable if there is a constant $M \in \mathbb{R}$ independent of $\mathcal{H}_T$, such that $|\tilde{v}^m_\alpha| \le M$ for all $(m, \alpha) \in \overline{\mathcal{H}_T}$.*

The truncation error of the scheme is closely connected with its consistency. The scheme is consistent in the interior of the domain $\mathscr{D}_T$ if and only if the truncation error converges to 0 as $|\mathcal{H}_T| \to 0$, and there is often a relation between the order of the truncation error and the rate of convergence of the scheme.

Definition 12.4 (Truncation error). *The truncation error of the finite difference operator* (11.8) *is defined by*

$$\tau^m_{\mathcal{H}_T,\alpha}(\phi) = \sup_{a\in\mathcal{A}} S^a_{\mathcal{H}_T}(t_m, X_\alpha, \phi) - \sup_{a\in\mathcal{A}} S^a(t_m, X_\alpha, \phi) \quad (12.4)$$

for $\phi \in C^\infty(\overline{\mathscr{D}_T})$ and $(m, \alpha) \in \mathcal{G}_T$.

Before we prove consistency, we will find an expression for the truncation error of our scheme.

Lemma 12.5. *The truncation error $\tau^m_{\mathcal{H}_T,\alpha}(\phi)$ of the finite difference operator* (11.8) *satisfies*

$$\begin{aligned} \tau^m_{\mathcal{H}_T,\alpha}(\phi) \le K\Big(&\Delta t\big|\partial_t^2\phi\big|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x\big|\partial_x\phi\big|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x\big|\partial_x^2\phi\big|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x\big|\partial_x^3\phi\big|_{L^\infty(\overline{\mathcal{D}_T})} \\ &+ \Delta x\big|\partial_x^4\phi\big|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta y\big|\partial_y^2\phi\big|_{L^\infty(\overline{\mathcal{D}_T})}\Big) \end{aligned} \quad (12.5)$$

for some constant $K \in \mathbb{R}$ and all $\phi \in C^\infty(\overline{\mathcal{D}_T})$, where K is independent of $\mathcal{H}_T$ and ϕ, but dependent of $\mathcal{D}_T$.

Proof. It is shown in [12] that

$$\mathcal{I}^{a,\pm}[\phi](t,X) = \mathcal{I}^{a,\pm}_{\mathcal{H}_T}[\phi](t,X) + E^{\pm}_Q + E^{\pm}_{FDM} + E^{\pm}_I,$$

where E_Q, E_{FDM} and E_I denote, respectively, the error contributions from the approximation of the integral, the difference approximation and the second order interpolation, and can be estimated as follows

$$\begin{aligned}
|E^{\pm}_Q| &\leq \Delta x |\partial^3_z \phi(\cdot + \eta^a)|_{L^\infty(\overline{\mathcal{D}_T})} \int_{\mathbb{R}} \tilde{k}(z)\,dz \\
&\leq K\Delta x \Big(|\partial_x \phi(\mathcal{D}_T)|_{L^\infty(\overline{\mathcal{D}_T})} + |\partial^2_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + |\partial^3_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} \Big), \\
|E^{\pm}_{FDM}| &\leq \frac{1}{24} \Delta z^2 |\partial^4_z \phi(\cdot + \eta^\pi)|_{L^\infty(\overline{\mathcal{D}_T})} \int_{\mathbb{R}} \tilde{k}(z)\,dz \\
&\leq K\Delta x \Big(|\partial_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + |\partial^2_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + |\partial^3_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + |\partial^4_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} \Big),
\end{aligned}$$

and

$$\begin{aligned}
|E^{\pm}_I| &\leq 4 \frac{\Delta x^2}{\Delta z^2} |\partial_x \phi(\cdot + \eta^\pi)|_{L^\infty(\overline{\mathcal{D}_T})} \int_{\mathbb{R}} \tilde{k}(z)\,dz \\
&= 4\Delta x |\partial_x \phi(\cdot + \eta^\pi)|_{L^\infty(\overline{\mathcal{D}_T})} \int_{\mathbb{R}} \tilde{k}(z)\,dz \\
&\leq K\Delta x \Big(|\partial_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + |\partial^2_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} \Big).
\end{aligned}$$

The term $\tilde{b}^a \delta^-_{x,\Delta x}\phi$ is also consistent, and by Taylor expansions, we see that

$$|\tilde{b}^a \partial_x \phi - \tilde{b}^a \delta^-_{x,\Delta x}\phi|_{L^\infty(\overline{\mathcal{D}_T})} \leq \frac{1}{2} \Delta x |\partial^2_x \phi|_{L^\infty(\overline{\mathcal{D}_T})}.$$

We see easily that $\mathcal{L}^a$ also is consistent, and Taylor expansions give us

$$\big|\mathcal{L}^a[\phi] - \mathcal{L}^a_{\mathcal{H}}[\phi]\big|_{L^\infty(\overline{\mathcal{D}_T})} \leq K\Big(|\partial^2_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} \Delta x + |\partial^2_y \phi|_{L^\infty(\overline{\mathcal{D}_T})} \Delta y \Big).$$

Doing a Taylor expansion of the time difference operator too, we get

$$\begin{aligned}
\sup_{(m,\alpha)\in\mathcal{G}_T} |\tau^m_{\mathcal{H}_T,\alpha}| &\leq \sup_{a\in\mathcal{A}} |S^a - S^a_{\mathcal{H}_T}|_{L^\infty(\overline{\mathcal{D}_T})} \\
&\leq \frac{1}{2}\Delta t |\partial^2_t \phi|_{L^\infty(\overline{\mathcal{D}_T})} + |\mathcal{L}^a[\phi] - \mathcal{L}^a_{\mathcal{H}}[\phi]|_{L^\infty(\overline{\mathcal{D}_T})} + |\mathcal{I}^a[\phi] - \mathcal{I}^a_{\mathcal{H}}[\phi]|_{L^\infty(\overline{\mathcal{D}_T})} \\
&\leq K\Big(\Delta t |\partial^2_t \phi|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x |\partial_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x |\partial^2_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} \\
&\quad + \Delta x |\partial^3_x \phi|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x |\partial^4_x \phi|_{L^\infty(\overline{\mathcal{D}_T})}| + \Delta y |\partial^2_y \phi|_{L^\infty(\overline{\mathcal{D}_T})} \Big),
\end{aligned}$$

so (12.5) holds. □

Before we proceed to the proofs of consistency, monotonicity and stability, we will discuss shortly a connection between the truncation error and the expected rate of convergence of the scheme. First assume the HJB equation has a classical solution V_ϵ with bounded derivatives of all orders. Also assume that our numerical solution is identical to the exact solution at time step m, i.e., $\tilde{v}^m_{i,j} = V_\epsilon(t_m, x_i, y_j)$ for all $(i,j) \in \overline{\mathcal{G}}$. For simplicity, we only consider the purely explicit version of the scheme. We have

$$\tilde{v}_\alpha^{m-1} = \tilde{v}_\alpha^m + \Delta t f_\alpha^m + \sup_{a \in \mathcal{A}} \left(\Delta t \mathcal{L}_{\mathcal{H}}^a [\tilde{v}]_\alpha^m + \Delta t \mathcal{I}_{\mathcal{H}}^a [\tilde{v}]_\alpha^m \right)$$

for some $a \in \mathcal{A}$, which implies that

$$\tilde{v}_{i,j}^{m-1} - V_\epsilon(t_{m-1}, x_i, y_j) = \Delta t \tau_{\mathcal{H}_T,\alpha}^m [V_\epsilon].$$

The expression $\tilde{v}_{i,j}^{m-1} - V_\epsilon(t_{m-1}, x_i, y_j)$ is the error produced by the scheme at the current time step. We see from Lemma 12.5 that the error produced in one time step is of order

$$O\big(\Delta t(\Delta t + \Delta x + \Delta y)\big) = O(\Delta t^2 + \Delta t \Delta x + \Delta t \Delta y),$$

since V_ϵ and its derivatives of all orders are bounded. To calculate $\tilde{v}$ at a certain time, we need $O(\Delta t^{-1})$ time steps. The accumulated error in our approximation is therefore of order

$$O\big(\Delta t^{-1}(\Delta t^2 + \Delta t \Delta x + \Delta t \Delta y)\big) = O\big(\Delta t + \Delta x + \Delta y\big).$$

We expect the rate of convergence to be 1 in both Δx, Δy and Δt, and we see that the error of the scheme is of the same order as the truncation error.

Now assume the HJB equation has a classical solution, but that at least one of the derivatives of the solution are unbounded. This is the case for the exact solution formula derived in Section 7.1, as $V_x, V_y \to \infty$ when $x, y \to 0$. The truncation error is not of order $O(\Delta x + \Delta y + \Delta t)$ anymore, as the "constants" in front of Δx, Δy and Δt in (12.5) are non-existent. We may manage to determine the order of the truncation error by considering the order of which the derivatives converge to infinity, see Section 15. In general, we will obtain a lower rate of convergence than 1 if the derivatives of the solution are unbounded.

Now assume the HJB equation do not have a classical solution. In this case the above arguments do not work, as we cannot find the Taylor expansion of the value function. We have to use alternative techniques, and in general the rate of convergence may be lower than 1. In [12] the authors consider a problem that is identical to ours, except that it is defined on $[0, T] \times \mathbb{R}^n$, and that the HJB equation and the numerical scheme are written on a more general form. It is proved that there exists a constant $K \in \mathbb{R}$ such that

$$-K(\Delta t^{1/10} + \Delta x^{1/10} + \Delta y^{1/10}) \leq v(t_m, x_i, y_j) - \tilde{v}_{i,j}^m \leq K(\Delta t^{1/4} + \Delta x^{1/4} + \Delta y^{1/4}),$$

for all $(m, i, j) \in \mathcal{G}_T$, where v is a viscosity solution of the HJB equation such that v is Lipschitz continuous in X and Hölder continuous in t. With some additional assumptions on ν in a neighbourhood of 0, we can replace the terms $\Delta x^{1/4}$ and $\Delta y^{1/4}$ on the right hand side by $\Delta x^{1/2}$ and $\Delta y^{1/2}$, respectively. We can expect the error of our scheme to be of the same order, as our scheme is a special case of the scheme in [12], except for the domain of the value function, but we will not attempt to adapt the proof in [12] to our case.

Now we will prove that the scheme (11.21) is consistent, monotone and stable, provided Δt is sufficiently small. The exact condition that must be satisfied by Δt will be stated in Lemma 12.7.

Lemma 12.6. *The numerical scheme* (11.21) *is consistent.*

Proof. We see immediately that the scheme is consistent for interior points $(t,X) \in \mathscr{D}_T$, since the truncation error of the scheme converges to 0 when $|\mathcal{H}_T| \to 0$. We will therefore focus on proving consistency at the three boundaries $x = 0$, $y = 0$ and $t = T$. We start with the case $x = 0$, $y > 0$ and $t < T$, and we want to prove (12.1) and (12.2). For all (t_m, X_α) sufficiently close to (t,X), and ξ and $|\mathcal{H}_T|$ sufficiently small, we have

$$\mathcal{R}_{\mathcal{H}_T}(t,X,v) = \sup_{a\in\mathcal{A}} \mathcal{S}^a(t,X,v) \qquad \text{or} \qquad \mathcal{R}_{\mathcal{H}_T}(t,X,v) = g(t,X) - v(t,X).$$

By using techniques from the proof of Lemma 12.5, we see that

$$\lim_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n,X_\alpha)\to(t,X),\\ \xi\to 0}} \left(\sup_{a\in\mathcal{A}} \mathcal{S}^a_{\mathcal{H}_T}(t_n, X_\alpha, \phi + \xi) \right) = \sup_{a\in\mathcal{A}} \mathcal{S}^a(t,X,\phi),$$

and we see easily that

$$\lim_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n,X_\alpha)\to(t,X),\\ \xi\to 0}} g(t_n,X_\alpha) - \big(\phi(t_n,X_\alpha) + \xi\big) = g(t,X) - \phi(t,X),$$

so

$$\liminf_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n,X_\alpha)\to(t,X),\\ \xi\to 0}} \mathcal{R}_{\mathcal{H}_T}(t_n,X_\alpha,\phi) = \min\left\{ \sup_{a\in\mathcal{A}} S^a(t,X,\phi);\, g(t,X) - \phi(t,X) \right\} = \underline{\mathcal{R}}(t,X,\phi),$$

and (12.1) is proved. We can prove (12.2) by a similar argument. We can also prove (12.1) and (12.2) for $t = T$ by a similar argument. For the case $y = 0$, however, the proof is slightly different. If $y = 0$, we have

$$\underline{\mathcal{R}}(t,X,\phi) = \min\left\{ \sup_{a\in\mathcal{A}} S^a(t,X,\phi);\, p - \phi(t,X) \right\} = -\infty,$$

for all $x \geq 0$ and $t \in [0,T]$, so (12.1) obviously holds. If $y = 0$, $x > 0$ and $t < T$, we have

$$\begin{aligned}
\overline{\mathcal{R}}(t,X,\phi) &= \max\left\{ \sup_{a\in\mathcal{A}} S^a(t,X,\phi);\, p - \phi(t,X) \right\} \\
&= \sup_{a\in\mathcal{A}} S^a(t,X,\phi) \\
&= \limsup_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n,X_\alpha)\to(t,X),\\ \xi\to 0}} \mathcal{R}_{\mathcal{H}_T}(t_n,X_\alpha,\phi+\xi),
\end{aligned}$$

so (12.2) holds. We can easily generalize the above arguments to the points $(t,0,0)$, $(T,0,y)$ and $(T,x,0)$ for $t \in [0,T]$, $x \geq 0$ and $y \geq 0$. We see that (12.1) and (12.2) hold for all $(t,X) \in \overline{\mathscr{D}_T}$, and the lemma is proved. □

The following lemma says that our scheme is monotone provided Δt is sufficiently small.

Lemma 12.7 (Monotonicity). *The numerical scheme* (11.21) *is monotone if*

$$\sup_{a\in\mathcal{A}} \Delta t\Big((1-\theta)\bar{l}_{\tilde{\alpha}}^{a,m} + (1-\zeta)\bar{j}_{\tilde{\alpha}}^{a,m}\Big) \le 1, \tag{12.6}$$

for all $(m,\tilde{\alpha}) \in \overline{\mathcal{G}_T}$, *where* $\bar{l}_{\tilde{\alpha}}^{a,m}$ *and* $\bar{j}_{\tilde{\alpha}}^{a,m}$ *are positive constants of order* $O(\Delta x^{-2}+\Delta y^{-1})$ *defined in the proof below.*

Proof. We see immediately that appropriate functions $b_{\tilde{\alpha},\alpha}^{k,m}$ exist for $i=0$ and $j=y_{max}$, and that the case $i=x_{max}$ is a special case of $\tilde{\alpha}\in\mathcal{G}^*$, so we will assume $\tilde{\alpha}\in\mathcal{G}^*$. It is shown in [12] that $\mathcal{I}_{\mathcal{H}_T}^a$ is monotone, i.e., $\mathcal{I}_{\mathcal{H}_T}^a$ can be written as

$$\mathcal{I}_{\mathcal{H}_T}^a[\phi](t_m,X_{\tilde{\alpha}}) = \sum_{\alpha\in\overline{\mathcal{G}}} j_{\mathcal{H}_T,\alpha,\tilde{\alpha}}^{a,m}\Big(\phi(t_m,X_\alpha)-\phi(t_m,X_{\tilde{\alpha}})\Big),$$

where $j_{\mathcal{H}_T,\alpha,\tilde{\alpha}}^{a,m}\ge 0$. It is also shown that

$$\bar{j}_{\tilde{\alpha}}^{a,m} := \sum_{\alpha\in\overline{\mathcal{G}}} j_{\mathcal{H}_T,\alpha,\tilde{\alpha}}^{a,m} \le K_j\Delta x^{-1},$$

for some constant $K_j\in\mathbb{R}$. Since upwind differences are used when constructing $\mathcal{L}_{\mathcal{H}}^a$, we see that $\mathcal{L}_{\mathcal{H}}^a$ also is monotone, i.e.,

$$\mathcal{L}_{\mathcal{H}}^a[\phi](t_m,x_{\tilde{\alpha}}) = \sum_{\alpha\in\overline{\mathcal{G}}\setminus\{\tilde{\alpha}\}} l_{\mathcal{H}_T,\alpha,\tilde{\alpha}}^{a,m}\Big(\phi(t_m,X_\alpha)-\phi(t_m,X_{\tilde{\alpha}})\Big)$$

for $l_{\mathcal{H}_T,\alpha,\tilde{\alpha}}^{a,m}\ge 0$. We also see that

$$\bar{l}_{\tilde{\alpha}}^{a,m} := \sum_{\alpha\in\overline{\mathcal{G}}\setminus\{\tilde{\alpha}\}} l_{\mathcal{H}_T,\alpha,\tilde{\alpha}}^{a,m} \le K_l(\Delta x^{-2}+\Delta y^{-1})$$

for some $K_l\in\mathbb{R}$ independent of Δx and Δy, but dependent of x_{max} and y_{max}.

Writing (11.21) for $(n,\alpha)\in\mathcal{G}_T$ on the form (12.3), we get

$$b_{\tilde{\alpha},\tilde{\alpha}}^{n,m}(a) = \begin{cases} 1+\Delta t\theta\bar{l}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}^{a,m}+\Delta t\zeta\bar{j}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}^{a,m} & \text{for } m=n-1,\\ 1-\Delta t\Big((1-\theta)\bar{l}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}^{a,m}+(1-\zeta)\bar{j}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}^{a,m}\Big) & \text{for } m=n,\end{cases}$$

$$b_{\tilde{\alpha},\alpha}^{n,m}(a) = \begin{cases} \Delta t\theta l_{\mathcal{H}_T,\tilde{\alpha},\alpha}^{a,m}+\Delta t\zeta j_{\mathcal{H}_T,\tilde{\alpha},\alpha}^{a,m} & \text{for } m=n-1,\\ \Delta t(1-\theta)l_{\mathcal{H}_T,\tilde{\alpha},\alpha}^{a,m}+\Delta t(1-\zeta)j_{\mathcal{H}_T,\tilde{\alpha},\alpha}^{a,m} & \text{for } m=n,\end{cases}$$

and $b_{\tilde{\alpha},\alpha}^{n,m}\equiv 0$ for all other values of m. We must have $b_{\tilde{\alpha},\alpha}^{n,m}(a)\ge 0$ for all $a\in\mathcal{A}$ for the scheme to be monotone, i.e., the scheme is monotone if (12.6) is satisfied. □

If the scheme is monotone, it will also be stable.

Lemma 12.8. *The numerical scheme* (11.21) *is stable provided* (12.6) *holds.*

Proof. We will show by induction on n that

$$\max_{\alpha\in\overline{\mathcal{G}}} \tilde{v}^n_\alpha \leq |\widehat{W}|_{L^\infty(\overline{\mathcal{D}_T})} + (T-t_n)\sup_a |f^a|_{L^\infty(\overline{\mathcal{D}_T})}, \tag{12.7}$$

which will imply that the scheme is stable, as $\widehat{W}$, f and t are bounded on $\mathcal{D}_T$. The induction hypothesis is true for $n = N_T$ by (11.21) and (11.22a). Assume (12.7) holds for n. We wish to show that is also holds for $n-1$.

Let $\tilde{\alpha} \in \mathcal{G}^*$, such that $\tilde{\alpha} = (i,j)$ for $i \neq 0$. Let $a^* \in \mathcal{A}$ be the value of a that maximizes (11.22b). We know that a^* exists, because $\mathcal{A}$ is compact. We have

$$\begin{aligned}
&\Big(1 + \Delta t\theta \bar{l}^{a^*,n-1}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}} + \Delta t\zeta \bar{l}^{a^*,n-1}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}\Big)\tilde{v}^{n-1}_{\tilde{\alpha}} \\
&\qquad = \Big(1 - \Delta t(1-\theta)\bar{l}^{a^*,n}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}} - \Delta t(1-\zeta)\bar{l}^{a^*,n}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}\Big)\tilde{v}^n_{\tilde{\alpha}} \\
&\qquad\quad + \Delta t \sum_{\alpha\in\overline{\mathcal{G}}\backslash\{\tilde{\alpha}\}} \Big(\theta l^{a^*,n-1}_{\mathcal{H}_T,\alpha,\tilde{\alpha}} + \zeta j^{a^*,n-1}_{\mathcal{H}_T,\alpha,\tilde{\alpha}}\Big)\tilde{v}^{n-1}_\alpha \\
&\qquad\quad + \Delta t \sum_{\alpha\in\overline{\mathcal{G}}} \Big((1-\theta) l^{a^*,n}_{\mathcal{H}_T,\alpha,\tilde{\alpha}} + (1-\zeta) j^{a^*,n}_{\mathcal{H}_T,\alpha,\tilde{\alpha}}\Big)\tilde{v}^n_\alpha + \Delta t f^n_{\tilde{\alpha}}.
\end{aligned}$$

All coefficients on the right-hand side are positive provided (12.6) holds. Therefore we get

$$\begin{aligned}
&\Big(1+\Delta t\theta \bar{l}^{a^*,n-1}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}} + \Delta t\zeta \bar{l}^{a^*,n-1}_{\mathcal{H}_T,\tilde{\alpha},\tilde{\alpha}}\Big)\tilde{v}^{n-1}_{\tilde{\alpha}} \\
&\quad \leq \max_{\alpha\in\overline{\mathcal{G}}} v^n_\alpha + \Delta t \sum_{\alpha\in\overline{\mathcal{G}}} \Big(\theta l^{a^*,n-1}_{\mathcal{H}_T,\alpha,\tilde{\alpha}} + \zeta l^{a^*,n-1}_{\mathcal{H}_T,\alpha,\tilde{\alpha}}\Big) \max_{\alpha\in\overline{\mathcal{G}}} \tilde{v}^{n-1}_\alpha + \Delta t \max_{\alpha\in\overline{\mathcal{G}}} f^n_{\tilde{\alpha}}
\end{aligned} \tag{12.8}$$

for all $\tilde{\alpha} \in \mathcal{G}^*$. This inequality also holds for $\tilde{\alpha} \in \overline{\mathcal{G}}\backslash\mathcal{G}^*$: For $j = N_y$, the result follows by exactly the same argument as above, because the scheme for $j = N_y$ is identical to the scheme for $(n-1,\tilde{\alpha}) \in \mathcal{G}^*$, except for the restriction $c = 0$. For $i = N_x$, we see easily from (11.21) that

$$\tilde{v}^{n-1}_{N_x,j} \leq \max\big\{\tilde{v}^{n-1}_{N_x-1,j};\, \tilde{v}^{n-1}_{N_x,j+1}\big\},$$

so

$$\tilde{v}^{n-1}_{N_x,j} \leq \max_{\tilde{\alpha}\in\mathcal{G}^*} \tilde{v}^{n-1}_{\tilde{\alpha}}$$

by induction, and (12.8) holds. Combining the results above, we see that (12.8) holds for all $\tilde{\alpha} \in \overline{\mathcal{G}}$ with $i \neq 0$, and therefore

$$\max_{\alpha\in\overline{\mathcal{G}}, i\neq 0} \tilde{v}^{n-1}_\alpha \leq \max_{\alpha\in\overline{\mathcal{G}}} v^n_\alpha + \Delta t \max_{\alpha\in\overline{\mathcal{G}}} f^n_{\tilde{\alpha}}.$$

By the induction hypothesis for n, we get

$$\begin{aligned}
\max_{\substack{\alpha\in\overline{\mathcal{G}},\\ i\neq 0}} \tilde{v}^{n-1}_{\tilde{\alpha}} &\leq |\widehat{W}|_{L^\infty(\overline{\mathcal{D}_T})} + (T-t_n)\max_{\substack{\alpha\in\overline{\mathcal{G}}\\ m\geq n}} |f^m_\alpha|_{L^\infty(\overline{\mathcal{D}_T})} + \max_{\alpha\in\overline{\mathcal{G}}} \Delta t f^n_{\tilde{\alpha}} \\
&\leq |\widehat{W}|_{L^\infty(\overline{\mathcal{D}_T})} + (T-t_{n-1})\max_{\alpha\in\overline{\mathcal{G}}} |f^n_\alpha|_{L^\infty(\overline{\mathcal{D}_T})}.
\end{aligned}$$

For $i = 0$, we see from (11.21) that

$$\tilde{v}^{n-1}_{0,j} \leq |\widehat{W}|_{L^\infty(\overline{\mathcal{D}_T})} + (T-t_{n-1})\max_{\alpha\in\overline{\mathcal{G}}} |f^n_\alpha|_{L^\infty(\overline{\mathcal{D}_T})}.$$

Our claim (12.7) follows by induction.

Lemmas 12.6-12.8 imply that the numerical solution converges to the viscosity solution, but before we state this in a theorem, we will prove that (11.21) has a unique well-defined solution.

In Section 11.4 there is given a short sketch of the algorithm for finding $\tilde{v}$. The idea is to determine $\tilde{v}^{N_t}$ first by the terminal condition, and then go backwards in time, using $\tilde{v}^n$ to find $\tilde{v}^{n-1}$. We want to show that $\tilde{v}^{n-1}$ can be determined uniquely from $\tilde{v}^n$ at each time step.

The result is obvious in the case of an explicit scheme, as $\tilde{v}_\alpha^{n-1}$ can be expressed as an equation of $\tilde{v}^n$ for each $\alpha \in \overline{G}$. The result is harder to prove for the general case where θ and ζ can take any value in $[0,1]$.

When solving linear PDEs, a numerical scheme can be written as a linear system of equations $A\tilde{v} = b$. The problem of the scheme's solvability is reduced to determining whether the matrix A is singular or not. Our problem is not linear, because of the control a. If we fix a, we see easily that the scheme (11.21) is solvable: The matrix A corresponding to the scheme has $b_{\tilde{\alpha},\tilde{\alpha}}^{n,n-1}$ on the diagonals, and $b_{\tilde{\alpha},\alpha}^{n,n-1}$, $\tilde{\alpha} \neq \alpha$, in the other positions of the matrix. We see that the matrix is diagonally dominant, with at least some rows strictly diagonally dominant. It follows that A is non-singular.

However, this argument is not sufficient to prove that our scheme has a unique solution, because the control makes the system non-linear. We need to use another method to prove existence and uniqueness of the discrete equations. The following theorem uses the Banach fixed point theorem, see Theorem 2.29. The proof is relatively similar a proof in [12]. The main difference between the proofs is that we will take the boundary conditions of the scheme into consideration. We also need to do some modifications in the formulation because our problem is a terminal value problem, and not an initial value problem.

Lemma 12.9. *The numerical scheme* (11.21) *has a unique solution.*

Proof. As mentioned above, it is sufficient to show that the system of equations (11.21), (11.22b)-(11.22e) has a unique solution at each time step. Assume $\tilde{v}^n$ is given, and that we want to determine $\tilde{v}^{n-1}$. Let $\mathbb{R}^{\overline{\mathcal{H}}}$ denote the space of all real valued functions defined on $\overline{\mathcal{H}}$. Define the operator $T : \mathbb{R}^{\overline{\mathcal{H}}} \to \mathbb{R}^{\overline{\mathcal{H}}}$ by

$$(Tw)_\alpha = w_\alpha - \epsilon \mathcal{R}_{\mathcal{H}_T}(t_{n-1}, X_\alpha, \tilde{w}), \tag{12.9}$$

for all $\alpha \in \overline{\mathcal{H}}$, where $\tilde{w} : \overline{\mathcal{H}_T} \to \mathbb{R}$ satisfies $\tilde{w}^n = \tilde{v}^n$ and $\tilde{w}^{m-1} = w$. Note that

$$\mathcal{R}_{\mathcal{H}_T}(t_{n-1}, X_\alpha, \tilde{w}) = \begin{cases} \begin{aligned} \sup_{a\in\mathcal{A}} \Big\{ & -b_{\alpha,\alpha}^{n,n-1}(a) w_\alpha + \textstyle\sum_{\tilde{\alpha}\neq\alpha} b_{\alpha,\tilde{\alpha}}^{n,n-1}(a) w_{\tilde{\alpha}} \\ & + \textstyle\sum_{\tilde{\alpha}} b_{\alpha,\tilde{\alpha}}^{n,n}(a) \tilde{v}_{\tilde{\alpha}}^n + \Delta t f_\alpha^{n,a} \Big\} \end{aligned} & \text{for } \alpha \in \mathcal{G}^*, \\ -\left(\frac{\beta}{\Delta y} + \frac{1}{\Delta x}\right) w_{i,j} + \frac{\beta}{\Delta y} w_{i,j+1} \frac{1}{\Delta x} w_{i-1,j} & \text{for } i = N_x, \\ \begin{aligned} \sup_{(\pi,0)\in\mathcal{A}} \Big\{ & -b_{\alpha,\alpha}^{n,n-1}(a) w_\alpha + \textstyle\sum_{\tilde{\alpha}\neq\alpha} b_{\alpha,\tilde{\alpha}}^{n,n-1}(a) w_{\tilde{\alpha}}, \\ & + \textstyle\sum_{\tilde{\alpha}} b_{\alpha,\tilde{\alpha}}^{n,n}(a) \tilde{v}_{\tilde{\alpha}}^n + \Delta t f_\alpha^{n,a} \Big\} \end{aligned} & \text{for } j = N_y, \\ -w_\alpha + g_j^n & \text{for } i = 0, \\ w_{N_x-1,N_y-1} - w_\alpha & \text{for } \alpha = (N_x, N_y). \end{cases}$$

First we will show that T is a contraction for sufficiently small ϵ, i.e.,

$$|Tw_\alpha - T\tilde{w}_\alpha|_{L^\infty} \leq (1-\epsilon)|w_\alpha - \tilde{w}_\alpha|_{L^\infty}.$$

Let $\epsilon > 0$ be so small that

$$\epsilon\big(1 + \Delta t(\theta \bar{l}_\alpha^{a,n} + \zeta \bar{j}_\alpha^{a,n})\big) \leq 1.$$

Using $b_{\tilde{\alpha},\alpha}^{n,m} \geq 0$, we see that

$$\begin{aligned}
&Tw_\alpha - T\tilde{w}_\alpha \\
&\quad \leq \sup_{a\in\mathcal{A}} \bigg\{ \Big(1 - \epsilon\big(1 + \Delta t(\theta \bar{l}_\alpha^{a,n} + \zeta \bar{j}_\alpha^{a,n})\big)\Big)(w_\alpha - \tilde{w}_\alpha) \\
&\qquad + \epsilon \Delta t(\theta \bar{l}_\alpha^{a,n} + \zeta \bar{j}_\alpha^{a,n})|w_{\tilde{\alpha}} - \tilde{w}_{\tilde{\alpha}}|_{L^\infty} \bigg\} \\
&\quad \leq (1-\epsilon)|w - \tilde{w}|_{L^\infty},
\end{aligned}$$

for all $\alpha \in \mathcal{G}^*$. Interchanging the role of $\tilde{w}$ and w, we see that

$$|Tw_\alpha - T\tilde{w}_\alpha| \leq (1-\epsilon)|w - \tilde{w}|_{L^\infty(\overline{\mathcal{G}})} \tag{12.10}$$

for all $\alpha \in \mathcal{G}^*$. We see easily that (12.10) also is satisfied for $\alpha \in \overline{\mathcal{G}}\backslash\mathcal{G}^*$. Taking the supremum over all $\alpha \in \overline{\mathcal{G}}$, we see that T is a contraction under the supremum norm. Using the Banach fixed point theorem 2.29, we see that T has a unique fixed point w^*, and we see by insertion in 12.9 that w^* solves (11.21). $\square$

The theorem above does not only prove that the system of equations (11.21), (11.22b)-(11.22e) can be uniquely solved. It also gives us an algorithm for determining $\tilde{v}^{n-1}$ at each time step, as there is an iterative procedure associated with the Banach fixed point theorem, see Theorem 2.29.

To prove that $\tilde{v}$ converges to V_ϵ, we need the following strong comparison principle.

Theorem 12.10 (Strong Comparison Principle). *If $\underline{v}$ is a subsolution of* (11.5) *on $\overline{\mathcal{D}_T}$, and $\overline{v}$ is a supersolution of* (11.5) *on $\overline{\mathcal{D}_T}$, we have $\underline{v} \leq \overline{v}$ on $\overline{\mathcal{D}_T}$.*

The following theorem is the main result of the section. The proof is almost identical to a similar proof in [5]. However, we still choose to include a sketch of the proof, as the problem in [5] is formulated as a local HJB equation, not an integro-PDE. The theorem is based on the the strong comparison principle of Theorem 12.10, in addition to Lemmas 12.6-12.8.

Theorem 12.11. *The numerical scheme* (11.21) *converges locally uniformly to the viscosity solution of* (8.3) *and* (8.5)*, provided the strong comparison principle of Theorem 12.10 is satisfied.*

Proof. Let $\overline{v}, \underline{v} : \overline{\mathcal{D}_T} \to \mathbb{R}$ be defined by

$$\overline{v}(t,X) = \liminf_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n,X_\alpha)\to(t,X)}} \tilde{v}_\alpha^n$$

and

$$\underline{v}(t,X) = \limsup_{\substack{|\mathcal{H}_T|\to 0,\\ (t_n,X_\alpha)\to(t,X)}} \tilde{v}^n_\alpha.$$

We will prove that $\overline{v}$ is a supersolution of (8.3) and (8.5), and that $\underline{v}$ is a subsolution of (8.3) and (8.5). Suppose $(t,X) \in \overline{\mathcal{D}_T}$ is a global maximum of $\underline{v} - \phi$ for some $\phi \in C^{1,2,1}(\overline{\mathcal{D}_T})$. By a similar argument as in the proof of Lemma 6.5, we may assume the maximum is strict, and that $\underline{v}(t,X) = \phi(t,X)$. Let $\{\mathcal{H}_{T,k}\}_{k\in\mathbb{N}}$ be a sequence of grids, such that $|\mathcal{H}_{T,k}| \to 0$ as $k \to \infty$, and let $\tilde{v}_k$ denote the numerical solution corresponding to the grid $\mathcal{H}_{T,k}$. Let (t_k, X_k) be a global maximum of $\tilde{v}_k - \phi$. By using a similar technique as in the proof of Lemma 10.28, we see that $(t_k, X_k) \to (t,X)$ and $\tilde{v}_k(t_k, X_k) \to \underline{v}(t,X)$ as $k \to \infty$.

By the monotonicity of $\mathcal{R}_{\mathcal{H}_{T,k}}$ and the definition of $\tilde{v}_k$, we have

$$\mathcal{R}_{\mathcal{H}_{T,k}}(t_k, X_k, \phi + \xi_k) \geq \mathcal{R}_{\mathcal{H}_{T,k}}(t_k, X_k, \tilde{v}_k) = 0.$$

Taking limits and using the consistency of $\mathcal{R}_{\mathcal{H}_{T,k}}$, we get

$$\begin{aligned} 0 &\leq \limsup_{k\to\infty} \mathcal{R}_{\mathcal{H}_{T,k}}(t_k, X_k, \phi + \xi_k) \\ &\leq \limsup_{\substack{|\mathcal{H}_{T,k}|\to 0,\\ (t_n,X_\alpha)\to(t,X)}} \mathcal{R}_{\mathcal{H}_T}(t, X, \phi + \xi) \\ &\leq \overline{\mathcal{R}}_{\mathcal{H}_T}(t, X, \phi), \end{aligned}$$

and it follows that $\underline{v}$ is a subsolution of (11.5). We can prove by a similar argument that $\overline{v}$ is a viscosity supersolution of (11.5).

The strong comparison principle of Theorem 12.10 implies that $\overline{v} \geq \underline{v}$. On the other hand, the definition of $\overline{v}$ and $\underline{v}$ gives us $\overline{v} \leq \underline{v}$. It follows that $\underline{v} = \overline{v}$, and that

$$v := \underline{v} = \overline{v} = \lim_{k\to\infty} \tilde{v}_k$$

is a viscosity solution of (11.5). By Lemma 11.2 and uniqueness results for viscosity solutions, v is the unique viscosity solution of (8.3) and (8.5). The convergence is locally uniform by the definitions of $\overline{v}$ and $\underline{v}$. □

Chapter 13
Scheme for the one-dimensional problem

In section 5.1 it was proved that the dimension of the HJB equation can be reduced if U and W satisfy (5.14). Reduction of the dimension of the problem is a huge advantage when constructing numerical schemes, as the running time and memory requirements become smaller. We will first derive a penalty approximation for the one-dimensional HJB equation, and then construct a numerical scheme for the penalty approximation. We will not provide proofs of convergence, only give heuristic arguments of why we have convergence. However, simulations in Chapter 15 will show that our numerical solution do converge to the analytical solution.

We proved in Chapter 10 that the solution of the penalty problem (8.3) and (8.5) converges uniformly to the value function V of the original problem. We do not have a similar result for the problem in reduced dimension. In fact, V_ϵ cannot be written as a function of only t and r ($r = x/y$, $r = x/(x+y)$ or $r = y/x$), so there is no one-dimensional equation in t and r that corresponds to (8.3).

We can show this by contradiction. Assume $V_\epsilon(t,x,y) = \overline{V}_\epsilon(t,r)$, where $r = x/y$. Introducing the short hand notations $F = F\left(t, X, D_X V, D_X^2 V, \mathscr{I}^\pi(t,X,V)\right)$, $\overline{F} = \overline{F}\left(t, r, \partial_r \overline{V}, \partial_r^2 \overline{V}, \mathscr{I}^\pi(t,r,\overline{V})\right)$, $G = G(D_X V)$ and $\overline{G} = \overline{G}(r, \overline{V}, \partial_r \overline{V})$, we know from Section 5.1 that $F = y^\gamma \overline{F}$ and $G = y^{\gamma-1}\overline{G}$. The value function V_ϵ satisfies $F + \max\{G; 0\}/\epsilon = 0$, so $\overline{V}_\epsilon$ must satisfy $y^\gamma \overline{F} + y^{\gamma-1}\max\{\overline{G}; 0\}/\epsilon = 0$ or, equivalently,

$$\overline{F}\left(t, r, \partial_r \overline{V}, \partial_r^2 \overline{V}, \mathscr{I}^\pi(t,r,\overline{V})\right) + \frac{1}{y\epsilon}\max\left\{\overline{G}(r, \overline{V}, \partial_r \overline{V}); 0\right\}.$$

We have obtained a contradiction, since this equation contains the variable y in addition to t and r. We get similar results for $r = x/(x+y)$ and $r = y/x$.

The numerical scheme we constructed in Chapter 11, was based on the penalty approximation (8.3) of (5.11). There is no one-dimensional equivalent of this penalty approximation, so we must find another penalty approximation if we wish to solve a one-dimensional system. From the form of (8.3), it is natural to try this penalty approximation:

$$\overline{F} + \frac{1}{\epsilon}\max\{\overline{G}; 0\} = 0. \tag{13.1}$$

The functions $\overline{F}$ and $\overline{G}$ are defined by (5.17) and (5.16), (5.21) and (5.20), and (5.23) and (5.22), respectively, for the three cases $r = x/y$, $r = x/(x+y)$ and $r = y/x$. Note that the penalty approximation (13.1) corresponds to a control problem where the the control is absolutely continuous, and the consumption rate is bounded by y/ϵ, $(x+y)/\epsilon$

and x/ϵ, respectively, for the three cases $r = x/y$, $r = x/(x+y)$ and $r = y/x$. We do not have any convergence results for (13.1), but it seems reasonable that the solution converges to $\overline{V}$. The heuristic proofs of convergence given in Section 10.1 also apply to the penalty approximation (13.1).

We try to construct a stable, consistent and monotone finite difference scheme for the penalty approximation (13.1). We will see that it is difficult to prove stability in the case $r = x/y$, while handling the integral term is challenging for the cases $r = x/(x+y)$ and $x = y/x$.

We start with the case $r = x/y$. Replacing all derivatives in (13.1) by upwind differences, the scheme can be written as

$$\sup_{a\in\mathcal{A}} \overline{\mathcal{S}}^a_{\mathcal{H}'_T}(t_n, X_\alpha, \tilde{v}) = 0, \tag{13.2}$$

where

$$\begin{aligned}\overline{\mathcal{S}}^a_{\mathcal{H}'_T}(t_n, r_i, \tilde{v}) = &\frac{1}{\Delta t}\left(\tilde{v}^n_i - \tilde{v}^{n-1}_i\right) + f^{n-1}_i + \theta\overline{\mathcal{L}}^a_{\mathcal{H}'_T}[\tilde{v}]^{n-1}_i + (1-\theta)\overline{\mathcal{L}}^a_{\mathcal{H}'_T}[\tilde{v}]^n_i \\ &+ \zeta\overline{\mathcal{I}}^a_{\mathcal{H}'_T}[\tilde{v}]^{n-1}_i + (1-\zeta)\overline{\mathcal{I}}^a_{\mathcal{H}'_T}[\tilde{v}]^n_i\end{aligned} \tag{13.3}$$

for some $\theta, \zeta \in [0,1]$,

$$f^n_i = e^{-\delta t_n}U(1),$$

$$\begin{aligned}\overline{\mathcal{L}}^a_{\mathcal{H}'_T}[\tilde{v}] = &\frac{1}{2}(\sigma\pi r)^2\Delta_{rr,\Delta r}\tilde{v} + \big(\hat{r} + (\hat{\mu} - \hat{r})\pi\big)r\delta^+_{r,\Delta r}\tilde{v} + \beta r\delta^+_{r,\Delta r}\tilde{v} - \beta\gamma\tilde{v} \\ &+ c\Big(-\beta\delta^-_{r,\Delta r}\tilde{v} + \beta\gamma\tilde{v} - \delta^-_{r,\Delta r}\tilde{v}\Big),\end{aligned}$$

the discretization $\overline{\mathcal{I}}_{\mathcal{H}'_T}$ of $\mathscr{I}^\pi$ similar to $\mathcal{I}_{\mathcal{H}}$, and

$$\mathcal{H}'_T = \big\{(t,r) \in \mathbb{R}^2 : t = m\Delta t,\ r = i\Delta r,\ m \in \{1,\dots,N_t\},\ i \in \{1,\dots,N_r\}\big\}$$

for $\Delta t = T/N_t$, $\Delta r = R_{max}/N_r$, $N_t, N_r \in \mathbb{N}$. By (5.18), we let the boundary condition for $r = 0$ be

$$\tilde{v}^m_0 = \int_t^T U\left(e^{-\beta(s-t)}\right) ds + \widehat{W}\left(0, e^{-\beta(T-t)}\right) \tag{13.4}$$

for $m = 0, 1, \dots, N_t$, and by (5.19) we let the terminal condition of the scheme be

$$\tilde{v}^{N_t}_i = \widehat{W}(r_i, 1)$$

for $i = 0, 1, \dots, N_r$. We employ a Neumann boundary condition at $r = R_{max}$, i.e.,

$$\tilde{v}^m_{N_r} = \tilde{v}^m_{N_r-1}$$

for $m = 0, 1, \dots, N_t - 1$.

In order to get a monotone scheme, the coefficients of all $\tilde{v}^n_{i'}$ in (13.3) must be positive. This is automatically satisfied for $i' = i \pm 1$, since we employed upwind differences. It is also satisfied for $i' = i$ for sufficiently small Δt.

Proceeding as in the proof Lemma 12.8, we get

$$\max_{0\le i\le N_r} \tilde{v}^{m-1}_i \le \left(1 + \frac{\beta\gamma}{\epsilon}\Delta t - \beta\gamma\Delta t\right) \max_{0\le i\le N_r} \tilde{v}^m_i + \max_{0\le i\le N_r} f^m_i.$$

Since f is bounded, and ϵ and Δt are assumed to be small, we see that

$$\max_{0\le i\le N_r} \tilde{v}_i^{m-1} \lesssim \left(1+\frac{\beta\gamma}{\epsilon}\Delta t\right) \max_{0\le i\le N_r} \tilde{v}_i^m,$$

so

$$\max_{0\le i\le N_r} v_i^0 \lesssim \left(1+\frac{\beta\gamma}{\epsilon}\Delta t\right)^{T/\Delta t} \max_{0\le i\le N_r} \tilde{v}_i^{N_t}.$$

We know that $(1+\beta\gamma\Delta t/\epsilon)^{T/\Delta t} \to e^{T\beta\gamma/\epsilon}$ as $\Delta t \to 0$. But $e^{T\beta\gamma/\epsilon} \to \infty$ when $\epsilon \to 0$, so we do not manage to establish a maximum principle.

The scheme may converge, even though we have not managed to prove a maximum principle. It was the term $c\beta\gamma\tilde{v}$ that caused the problem in the maximum principle proof, and in the above paragraphs, we assumed the worst-case situation $c = 1/\epsilon$ always will occur. In real simulations, we will probably have $c = 0$ in a large fraction of the iterations.

In fact, the following probabilistic induction argument in m and i shows that we expect the scheme to satisfy some kind of maximum principle for smooth value functions, because $c = 1/\epsilon$ only if $|v_{i-1}^m - v_i^m|$ is small, and because $\tilde{v}_0^m$ is bounded for all $m \in \{0,\dots,N_t\}$: We will assume $\theta,\zeta = 0$, i.e., the scheme is purely explicit, but the argument works equally well for $\theta,\zeta > 0$. Assume we have established an upper bound on $\tilde{v}^m$, and want to find an upper bound on $\tilde{v}^{m-1}$. We see by (13.4) that $\tilde{v}_0^{m-1}$ is bounded. Assume we have found some appropriate upper bound on $\tilde{v}_{i-1}^{m-1}$, and want to calculate $\tilde{v}_i^{m-1}$. If $c = 0$, we see immediately that $\tilde{v}_i^{m-1} \le \max_{0\le j\le N_r} \tilde{v}_j^m + f_i^m\Delta t$, so we have found an upper bound on $\tilde{v}_i^{m-1}$. If $c = 1/\epsilon$, we must have

$$-\beta\delta_{r,\Delta r}^{-}\tilde{v}_i^m + \beta\gamma\tilde{v}_i^m - \delta_{r,\Delta r}^{-}\tilde{v}_i^m \ge 0 \qquad \Leftrightarrow \qquad |\tilde{v}_i^m - \tilde{v}_{i-1}^m| \le \frac{\beta\gamma\Delta r}{\beta r_{i-1} + 1 - \gamma\Delta r}\tilde{v}_{i-1}^m,$$

since c is chosen to maximize $\overline{\mathbb{S}}^a_{\mathcal{H}\mathcal{C}'_T}$. For smooth value functions we expect $\tilde{v}_i^{m-1} - \tilde{v}_{i-1}^{m-1}$ to be of the same order as $\tilde{v}_i^m - \tilde{v}_{i-1}^m$, and therefore we expect to have $\tilde{v}_i^{m-1} - \tilde{v}_{i-1}^{m-1} = O(\Delta r)$. If we have found an upper bound on $\tilde{v}_{i-1}^{m-1}$, we may manage to find an upper bound also on $\tilde{v}_i^{m-1}$. We can find an upper bound on all elements of $\tilde{v}^{m-1}$ by induction.

For the case $r = y/x$, the situation is different than in the case above. Again the penalty approximation is given by (13.1), but $\overline{F}$ and $\overline{G}$ are given by (5.23) and (5.22), respectively. The coefficient in front of $\overline{V}$ in $\overline{G}$ is negative in this case, and therefore it is easy to establish a maximum principle. However, the discretization of the integral operator is more challenging in this case. In the case $r = x/y$ the integral operator is on the form described in [12], and we can use results from this article to find an appropriate discretization. In the case $r = y/x$ the form of the integral operator is different, and we would have to develop our own results.

For the case $r = x/(x+y)$ the properties of a discretization depends on the size of β. For $\beta \le 1$, the coefficient of $\overline{V}$ in $\overline{G}$ is negative, so we manage to establish a maximum principle. For $\beta > 1$, however, we will meet the same problem as described above for $r = x/y$. Just as in the case $r = y/x$, the integral term is on a different form than for the two-dimensional problem.

Only the scheme with $r = x/y$ will be implemented. The scheme will be used for simulations with jumps in the Lévy process. The time requirements for the integral calculations associated with the jumps become too large for the two-dimensional scheme described earlier.

Chapter 14
Implementation of the finite difference schemes

In this chapter we will explain and discuss some aspects of the implementation of the schemes described earlier. We will also discuss the time complexity of the schemes, and realistic values for the parameters of the schemes. We will focus on the explicit versions of the schemes from Chapters 11 and 13, as only these schemes will be implemented. The MATLAB implementation of the schemes can be found in the appendix.

Only the code for the one-dimensional problem admits jump processes. The reason for this, is that the calculation of the integral in the HJB equation is time consuming, and the time complexity of the scheme is reduced when the dimension is reduced.

As mentioned above, only the explicit versions of the schemes will be implemented. The advantage of a purely implicit scheme, is that fewer time steps are needed, as we have no constraint on the value of Δt. One advantage of a purely explicit scheme, is that each iteration is faster, as we do not have to solve a system of equations in each iteration. Let us assume the optimal control (π, c) is known, and that we want to employ an implicit numerical method. Since the optimal control is known, the system of equations we need to solve at each time step, is linear. If $\zeta > 0$, we have to solve a linear system of equations with full matrices in each iteration. This is in general a very time consuming operation. If only the other terms were treated implicitly $(\theta > 0, \zeta = 0)$, we had obtained a banded matrix, and systems with banded matrices are in general easier to solve than systems with full matrices, see [36]. However, also banded matrix systems take longer to solve than a purely explicit system, and if the integral term is treated explicitly, we will still have a constraint on Δt. As mentioned in Chapter 12, the Banach fixed point theorem provides a method for determining $\tilde{v}^{m-1}$ from $\tilde{v}^m$ at each time step, but the rate of convergence might be slow, such that we have to perform many matrix multiplications at each time step.

The programs are built up as sketched in Section 11.4: First we find v^{N_t} by using the terminal value condition, and then we iterate over the time steps. At each time step m we use (11.21)/(13.2) to determine $\tilde{v}^m$ at interior points, and afterwards we find the value of $\tilde{v}^m$ at the boundary.

The biggest challenge is to determine which controls we should use at each time step. It is easy to determine the value of c, as $\mathcal{S}^a_{\mathcal{H}_T}$ and $\overline{\mathcal{S}}^a_{\mathcal{H}'_T}$ are linear functions of c. It is also easy to find the optimal value of π if $\nu \equiv 0$. If $\nu \equiv 0$, $\mathcal{S}^a_{\mathcal{H}_T}$ and $\overline{\mathcal{S}}^a_{\mathcal{H}'_T}$ are second degree polynomials in π, and we can easily find the maximizing value of π. If the optimal value of π is larger than 1, we define $\pi = 1$, and if the optimal value of π is less than

0, we define $\pi = 0$. If we $\nu \not\equiv 0$, on the other hand, $\mathcal{S}^a_{\mathcal{H}'_T}$ and $\overline{\mathcal{S}}^a_{\mathcal{H}'_T}$ are highly non-linear functions of π, and it is more difficult to find a maximizing value of π.

We will find an approximately correct value of π in each iteration by assuming π is a uniformly continuous function of t. Let us assume we are considering the one-dimensional scheme (13.2). If we have found an optimal value π^* in time step m, we assume the optimal value of π in time step $m-1$ is in some interval $[\pi^* - \Delta\pi, \pi^* + \Delta\pi]$. We calculate $\overline{\mathcal{S}}_{\mathcal{H}'_T}$ for each of the three values $\pi^* - \Delta\pi$, π^* and $\pi^* + \Delta\pi$, and choose the value of π that makes $\overline{\mathcal{S}}_{\mathcal{H}'_T}$ largest. For the initial iteration $t = T$, it might be possible to calculate the optimal value of π analytically, as $V(T, x, y)$ is known. Otherwise, we can start out with a random value of π; the system will adjust the value of π over the first few time steps, such that it becomes optimal. We should not let $\Delta\pi$ be too large, as this may give us an inaccurate value of π. We should not let $\Delta\pi$ be too small either, as we need to have $\Delta\pi \geq \frac{\partial\pi}{\partial t}\Delta t$ to capture the changes in π over time. However, numerical simulation shows that $\tilde{v}$ is not very sensitive for the choice of π, and for the cases we simulate in the next chapter, π is approximately constant in time. Therefore this simple technique gives sufficiently good results.

As described in the next section, the calculated values of π and c are recorded in the simulation to provide the user with the simulated optimal controls. The number of π's and c's that are calculated by the program is often very large, as many time steps are needed. Therefore only a small fraction of the used values are saved at each time steps. A random number generator is used to determine which values that are recorded at each time step, but due to the large number of time steps, the result varies little between different runs with the same input parameters.

The most technically challenging part of the implementation is the implementation of the integral part of the HJB equation. Programs for two different kinds of Lévy measures are implemented, see equations (15.6) and (15.7). The integral part of the HJB equation is calculated in separate functions, called *calcJump_caseA* and *calcJump_caseB*, respectively. The functions take a vector of controls π and the discrete value function $\tilde{v}^m$ as input parameters, and return the calculated value of the integral for each r_i, $i \in \{1, \ldots, N_r - 1\}$. For each r_i, the calculated integral is a vector product between $\tilde{v}^m$ and a vector of weights. For each value of n, the contributions from each of the points $z_n - \Delta z$, z_n and $z_n + \Delta z$, $\Delta z = n\Delta r$, are added to the vector of weights. For large n, $z_n + \Delta z$ will correspond to grid points outside our grid, and these contributions to the integral term will simply be skipped. For this reason, the integral calculations will be inaccurate close to the boundary $r = R_{max}$. In the case where the Lévy measure is given by (15.6), we will also truncate the sum (11.10) in another way. For large values of n, the integral $\tilde{k}^{\pm}_{\mathcal{H},n}$ is very small, and therefore only the terms of (11.10) where n is small, will have any influence on the result.

We will implement the case $\nu(dz) = \lambda\delta_{-\xi}(dz)$, and compare the result with the explicit solution formula found in Section 7.1. We will have $\sigma \neq 0$, and therefore the solution formula (7.17) does not apply. To find the terminal value condition, we need to solve the system of equations (7.6) for ρ. It is not easy to solve this system of equations, and to overcome this problem, we let π^* be an input argument to the program in this case instead of δ. If π^* is given, it is easy to solve the first of the two equations for ρ numerically. As soon as ρ is determined, we use the second equation to find δ. However, note that π^* will *not* be used in the simulation, except to calculate δ and determining the initial value of π in the iterations.

The efficiency of the codes has been optimized in several ways. First, matrix operations are used as much as possible, as matrix operations go much faster than element-wise operations. Second, calculations of integrals required for jump processes have been calculated outside the loop iterating over time. Much time is saved by not having to perform numerical integration at each time step.

The time complexity of the scheme (11.21) with $\nu \equiv 0$ and $\theta, \zeta = 0$ is easily seen to be $O(N_x N_y N_t)$ if all operations are done element-wise. In order for the scheme to be monotone, we need to have $N_t = O(\frac{1}{\epsilon}N_y + \frac{1}{\epsilon}N_x + N_x^2)$, so if we are assuming the value of N_t is chosen as small as possible and all operations are done element-wise, the running time is of order

$$O\left(\frac{1}{\epsilon}N_x N_y^2 + \frac{1}{\epsilon}N_x^2 N_y + N_x^3 N_y\right)$$

for given N_x and N_y. However, the code has been optimized by using matrix operations instead of element-wise operations, and this affects the running time. The running time of matrix operations in MATLAB is not proportional to the size of the matrix for small matrices ($< 400^2$ elements), and therefore there is no easy relation between grid size and running time for our program. See Section 15.5 for a more comprehensive discussion of this.

The efficiency of the scheme (13.1) with $\nu \equiv 0$ is easily seen to be $O(N_r N_t)$ if all operations are done element-wise. We know that N_t must be of order $O\left(N_r^2 + \frac{1}{\epsilon}N_r\right)$ to get a monotone scheme. If we choose N_t as small as possible and perform all operations element-wise, the running time is of order

$$O\left(N_r^3 + \frac{1}{\epsilon}N_r^2\right).$$

However, since the program uses vector and matrix operations instead of element-wise operations, the real running time will be different. If there are jumps in the Lévy process, the calculation of the integral term will dominate the running time, as we do not manage to optimize the code with matrix operations in the integral calculation. If the Lévy measure is given by (15.7), the number of terms in (11.10) is proportional with N_r, so the running time is also expected to be linear in N_r. If the Lévy measure is given by (15.6), the connection between N_r and the number of terms is more complicated, due to the truncation algorithm described above, and the number of terms will depend on both N_r and b.

At the end of this chapter, we will discuss shortly realistic values of the parameters of the problem. The strict constraints on the parameters that are necessary in order to get a well-defined problem are given in Chapter 3; here we will only discuss which values are the most realistic, with reference to articles concerning similar problems. The impact and interpretation of the different parameters will be discussed in Section 15.2, where we see how varying the different parameters influence the simulation results.

We will interpret t as measured in years, and $\hat{r}$ and $\hat{\mu}$ is therefore the (expected) return of the safe and risky asset, respectively, during one year. Looking through articles concerning related problems ([17], [18], [31], [32]), we see that the value of $\hat{r}$ varies between 0.06 and 0.07, while the value of $\hat{\mu}$ varies between 0.05 and 0.3. We also note that the used value of $\hat{\mu}$ always is larger than the value of $\hat{r}$, as mentioned in Chapter 3.

The Lévy measure ν and the constant σ describe the variation in the return of the risky asset. Several realistic models used in the literature ([6], [21], [9]) let L_t be a pure-jump Lévy process, i.e., $\sigma = 0$. In the articles [17], [18], [31] and [32], where $\nu \equiv 0$, the value of σ varies between 0.2 and 0.4. The value of the discount factor δ varies between 0.1 and 7 in [17], and is set to 0.1 in [18].

Chapter 15
Simulation results

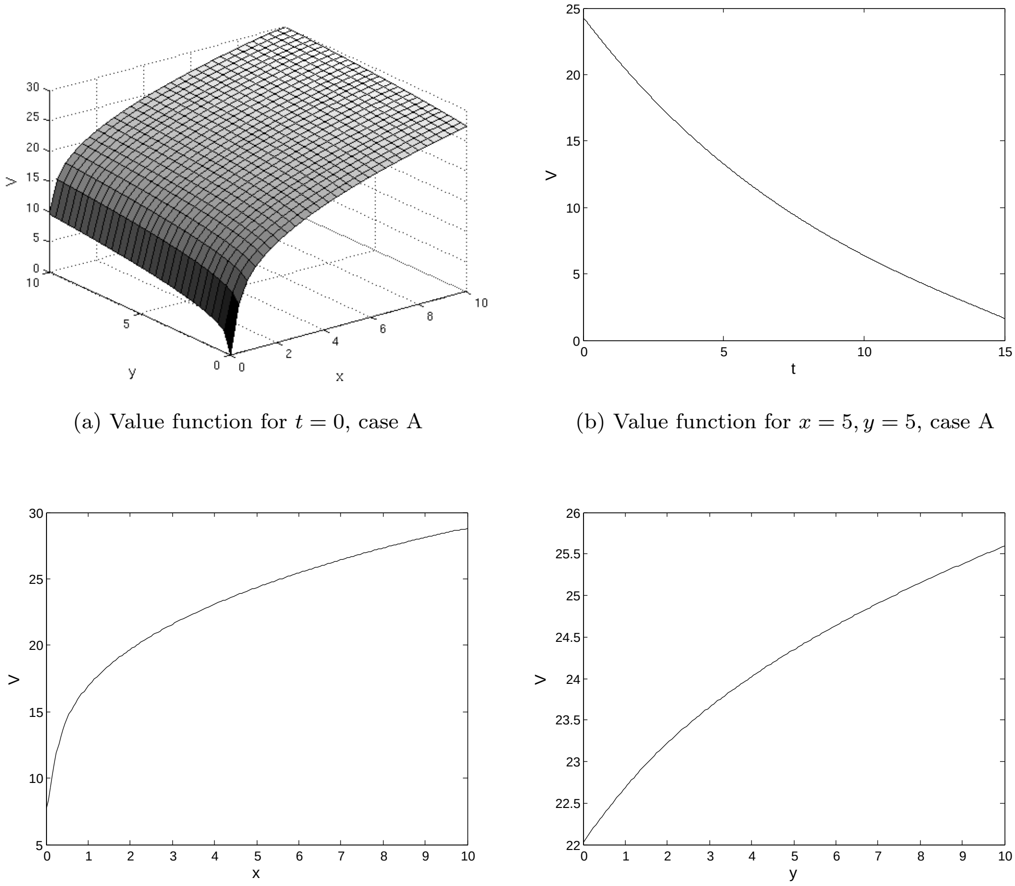
This section contains all the results from the simulations. In Sections 15.1-15.5 we will perform simulations with $\nu \equiv 0$, using the scheme described in Chapter 11, and in Sections 15.6-15.7, we will perform simulations with $\nu \not\equiv 0$, using the one-dimensional scheme described in Chapter 13. We describe typical features of the value function and the optimal controls found in the simulations, and we will see how the different parameters of the problem influence the result. We will also compare different boundary conditions, compare the singular problem to the non-singular problem, and discuss the rate of convergence and time complexity of the schemes.

We will have two reference cases, one which is employing a common choice of utility functions (A), and one which is based on the explicit solution formula found in Section 7.1 (B). We will consider three different kinds of plots:

1. plot of the value function V, in most cases for the time $t = 0$.
2. plot of the distribution of the value of π. The values of π that are calculated by the program for $G > 0$ are recorded, where G denotes a discrete approximation to $G(D_X V)$. The distribution of all used π-values are shown in a histogram.
3. plot of the relative frequency of $G > 0$ as a function of the ratio $x/(x + y)$.

The purpose of the π-plots is to find the optimal fraction of wealth to invest in the risky asset. We assume the investor will perform a consumption gulp if $G(D_X v) \geq 0$, and this implies that she will be in a state where $G(D_X V) < 0$ almost all the time. For this reason, only values of π where $G(D_X V) < 0$ are recorded. We are not recording the *used* values of π, which must be in the interval $[0, 1]$, but the value of $\pi \in (-\infty, \infty)$ that is optimal if we remove the constraint $\pi \in [0, 1]$. It is assumed that all values of π are in some interval $[a, b]$ $(a < 0, b > 1)$, and calculated values of π that are $< a$ and $> b$, are registered as a and b, respectively. Therefore the plot may show small frequency peaks at a and b, but these peaks are not real. For most π-plots we will have $a = 0$ and $b = 2$. We choose $b > 1$, because the calculated values of π often are above 1. If the optimal value of π is above 1, it is optimal for the agent to borrow money with rent $\hat{r}$ to invest in the risky asset. Our problem does not admit such behaviour, but it is still interesting to see the distribution of calculated π-values above 1.

The purpose of the G-plots is to see for which values of x, y and t the agent consumes. We know that $G(D_X V) < 0$ corresponds to consumption, because $G(D_X V) = \beta v_y - v_x < 0$ if $V(x - c, y + \beta c) < V(x, y)$ for all $c \in [0, x]$, i.e., if it is not optimal to consume. We see that $G(D_X V) \geq 0$ corresponds to consumption by a similar argument. For the



(a) Value function for $t = 0$, case A

(b) Value function for $x = 5, y = 5$, case A

(c) Value function for $y = 5, t = 0$, case A

(d) Value function for $x = 5, t = 0$, case A

Fig. 15.1: Plots of the value function for reference case A.

example studied in Section 7.1, the agent's consumption strategy is constant in time and depends only on the ratio x/y, but for other utility functions, the strategy may be different.

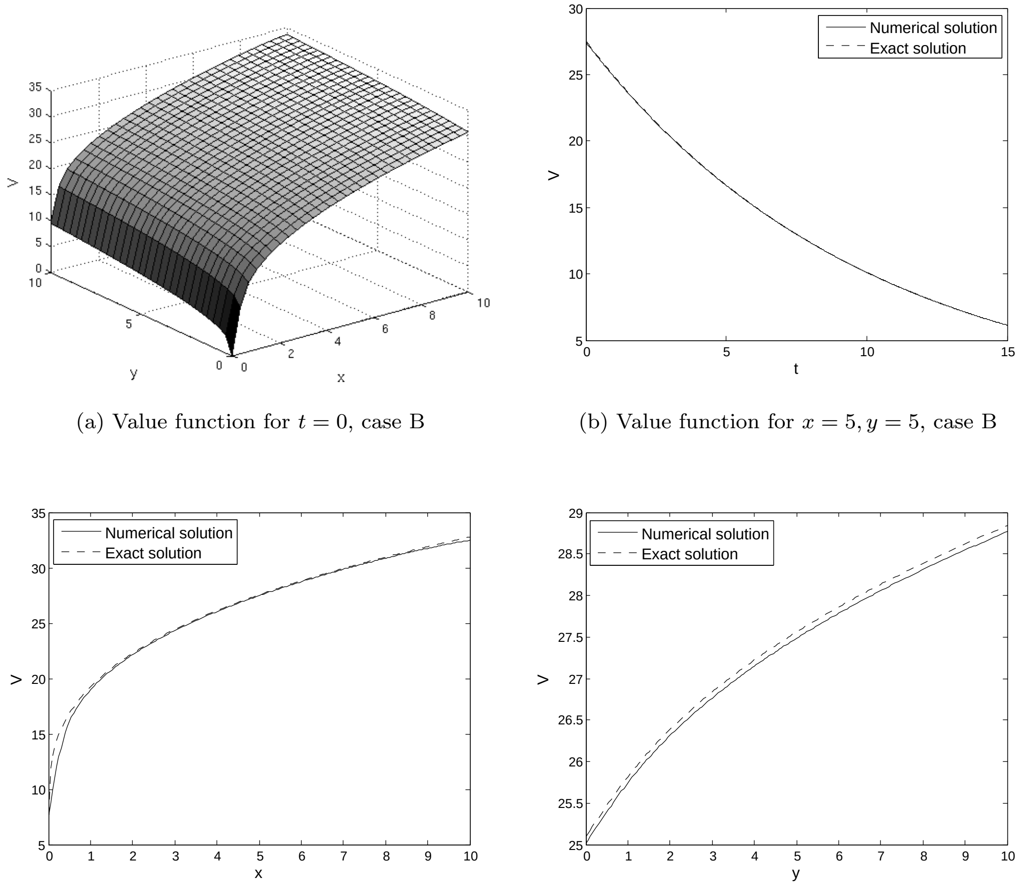
15.1 Reference cases

In this section we will define the reference cases A and B, and describe typical features of the value function and the optimal controls found in the simulations. Case B corresponds to the explicit solution found in Section 7.1, while case A corresponds to a more realistic choice of terminal utility function W.

In both case A and B, the utility function U is defined by

$$U(y) = \frac{y^\gamma}{\gamma}. \tag{15.1}$$

In case A the terminal utility function W is given by



(a) Value function for $t = 0$, case B

(b) Value function for $x = 5, y = 5$, case B

(c) Value function for $y = 5$, case B

(d) Value function for $x = 5$, case B

Fig. 15.2: Plots of the value function for reference case B, both the exact solution given by (7.27) and the simulated solution.

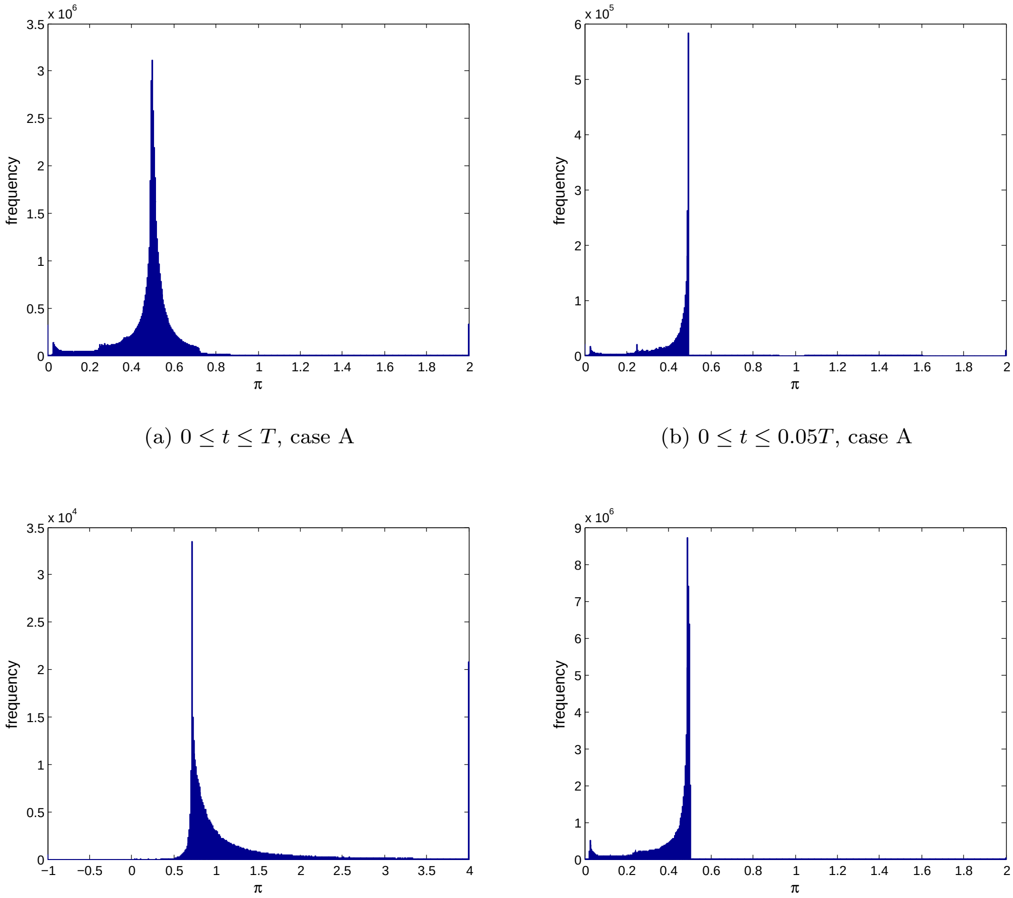
$$W(x, y) = \frac{De^{-\delta T}}{\gamma}(x + \beta y)^{\gamma}, \tag{15.2}$$

and in case B the terminal utility function W is given by (7.33). Note that (15.2) corresponds to a case where final satisfaction is independent of whether the agent does a final consumption gulp or not. The utility function (15.2) may also be interpreted as a case where satisfaction only depends on consumption (not final wealth), but where the investor always does a final consumption gulp at time T to get use of all the saved up wealth. Note that the function W defined by (15.2) satisfies

$$W(x, y) = \max_{c \in [0,x]} W(x - c, y + \beta y),$$

and therefore V_ϵ converges to V on the whole domain $\overline{\mathscr{D}_T}$, not only on $[0, T) \times \overline{\mathscr{D}}$, see Section 10.4.

Table 15.1 shows the reference values of all constants, in addition to the standard values of ϵ and $\mathcal{H}_T$. If other parameters are applied, it will be stated below the plot.



(a) $0 \le t \le T$, case A

(b) $0 \le t \le 0.05T$, case A

(c) $0.95T \le t \le T$, case B

(d) $0 \le t \le T$, case B

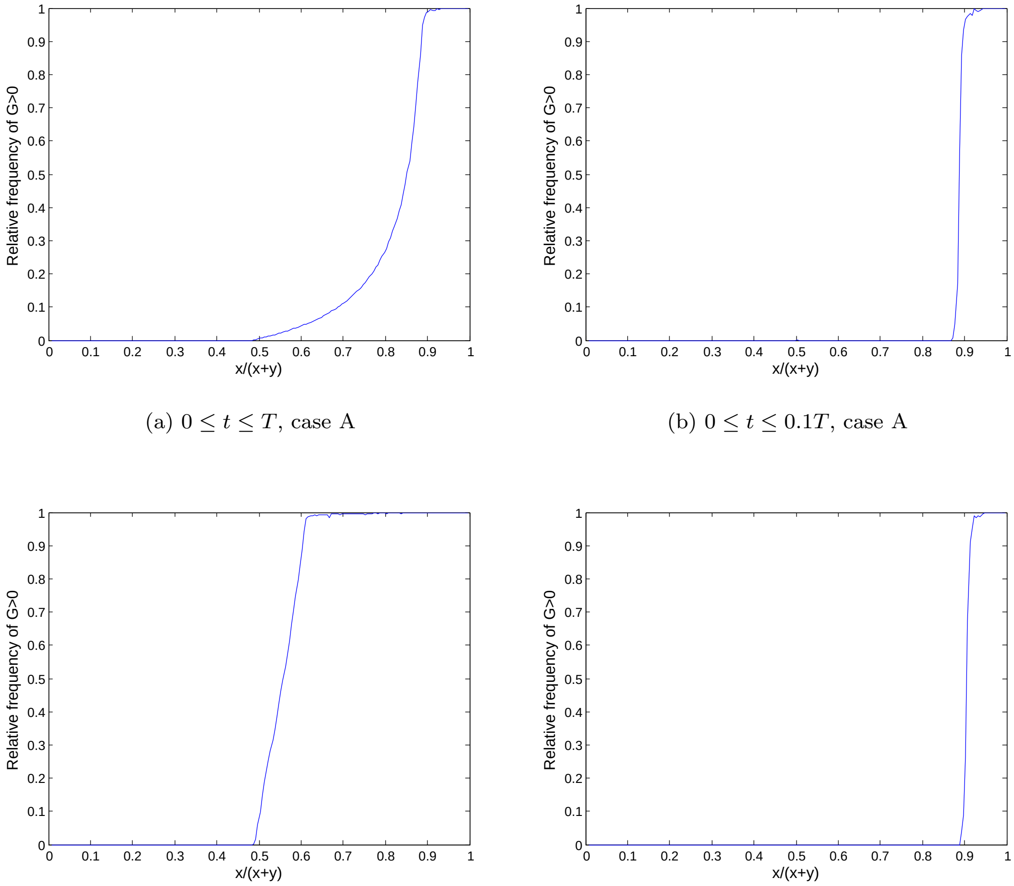
Fig. 15.3: Plot of the distribution of π. The spike around $\pi \approx 0.5$ on Figures 15.3b and 15.3d becomes more prominent if we decrease $|\mathcal{H}_T|$.

Case	T	δ	γ	D	σ	$\hat{r}$	$\hat{\mu}$	β	ν	U	W	Δx	Δy	ϵ
A	15	0.1	0.3	1	0.3	0.07	0.1	2	0	(15.1)	(15.2)	0.05	0.05	0.03
B	15	0.1	0.3	1	0.3	0.07	0.1	2	0	(15.1)	(7.33)	0.05	0.05	0.03

Table 15.1: This table describes the two reference cases A and B. The columns labelled U and W refer to equations.

See Figures 15.1 and 15.2 for plots of the value function in reference cases A and B. The value function is concave and increasing in x and y, as described in Lemmas 4.4 and 4.5. The value function decreases approximately exponentially with time, which corresponds well with the explicit solution formula (7.27) for case A. Exponential development in t is also reasonable from a financial point of view, as $X^{\pi,C}$ is expected in increase exponentially in the absence of consumption, while $Y^{\pi,C}$ will decrease exponentially in the absence of consumption.

See Figure 15.3 for plots of the distribution of π for reference cases A and B. In case B, $\pi \approx \pi^* \approx 0.503$ at all times and for all values of x and y, where π^* is defined as



(a) $0 \leq t \leq T$, case A

(b) $0 \leq t \leq 0.1T$, case A

(c) $0.9T \leq t \leq T$, case A

(d) $0 \leq t \leq T$, case B

Fig. 15.4: Plot of the sign of G as a function of $x/(x+y)$ for reference cases A and B. For the plots (b) and (c) only data from the time intervals $[0,\, 0.1T]$ and $[0.9T,\, T]$, respectively, are shown.

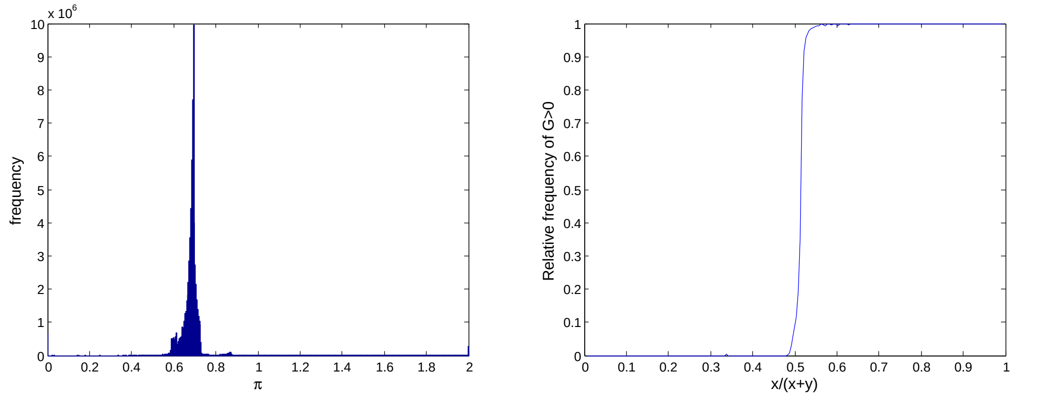
the solution of the system of equations (7.6). In case A, π increases with time, and is not constant in x and y. It is optimal to invest a large fraction of wealth in the safe asset initially, because too large early investments in the unsafe asset give too much uncertainty about the outcome. We see that the value of π converges to π^* when t decreases. When the investment horizon increases, the terminal value function becomes less important, and the solution converges to the solution in case B. See Lemma 4.7.

See Figure 15.4 for plots of the sign of G as a function of $x/(x+y)$. As explained above, $G > 0$ corresponds to consumption, while $G < 0$ corresponds to no consumption. In case B we have consumption if and only if $x/(x+y) \gtrsim k/(k+1) \approx 0.900$, where k is defined by (7.5), and the result is approximately constant in time. This corresponds well with the results in Section 7.1. In case A the consumption decreases with time. Again, we see that the solution in case A converges to the solution in case B when t decreases.

15.2 Varying parameter values

The purpose of this section is to find out how varying the different parameters of the problem influence the value function and the optimal investment strategy. We will only consider case A, as many of the results are similar for case A and case B.

Figure 15.5 shows how the value of δ influences the optimal choice of control. We see that the solution in case A converges faster to the solution in case B for a larger value of δ, as G and π are almost independent of time on Figure 15.5. The terminal value function becomes unimportant faster, because of the term $e^{-\delta T}$ in the definition of W, and therefore the solution converges faster. We see from (3.4) and (15.2) that the importance of early consumption increases, while the importance of late consumption and terminal wealth decreases, when we increase δ. Therefore the investor increases consumption, as we can see on Figure 15.5b.



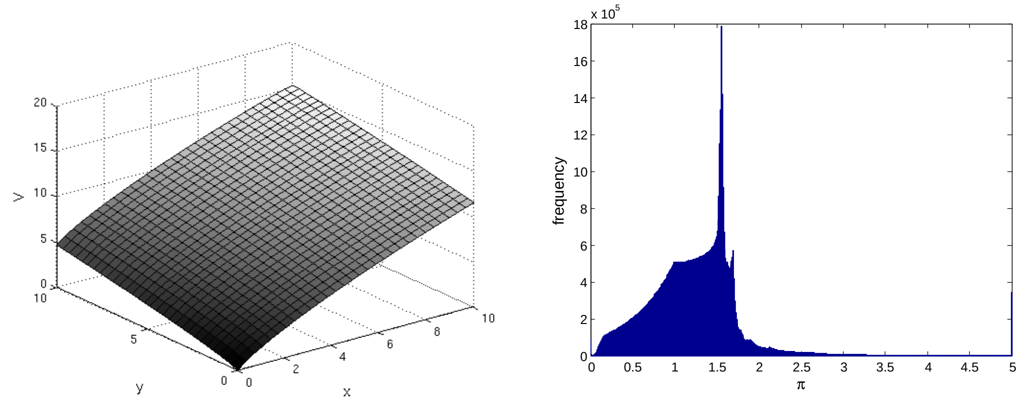
(a) The distribution of π for $0 \leq t \leq T$ (b) The sign of G as a function of $x/(x+y)$ for $0 \leq t \leq T$

Fig. 15.5: Plots for reference case A, where $\delta = 0.5$ instead of $\delta = 0.1$. Compare with Figures 15.3a and 15.4a.

Figure 15.6 shows how the value of γ influences our result. When γ is large, the utility functions are approximately linear, and hence the agent gives lower priority to safe investments, compared to investments with high expected return. Therefore the optimal value of π increases. We also see that the value function V is approximately linear for $\gamma = 0.8$, which corresponds well with the growth estimate given in Lemma 4.6.

Figure 15.7 shows how the value of D influences our result. When D is large, the investor gives high priority to final satisfaction, compared to satisfaction for $t < T$. Therefore the agent will decrease consumption, see Figure 15.7b. The value function is almost independent of y on Figure 15.7a. Since the agent consumes rarely, $Y_T^{\pi,C}$ will be small. It follows that the value of $X_T^{\pi,C}$ contributes the most to final satisfaction, and that V is almost independent of y.

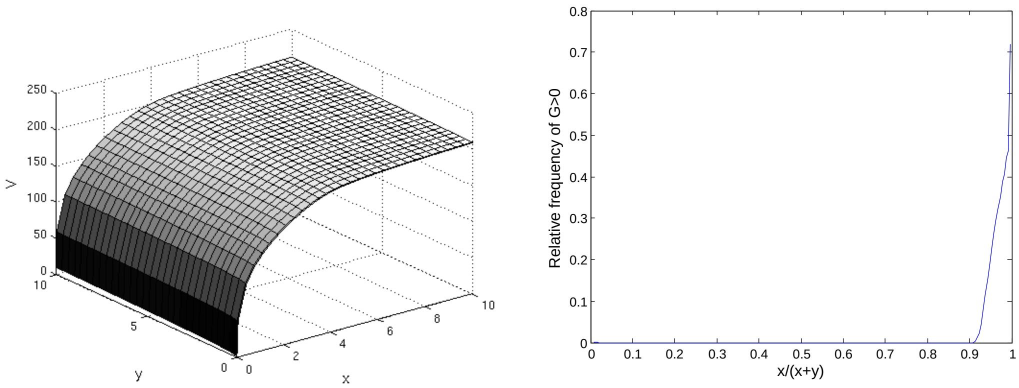
Figure 15.8 shows how the value of σ influences the optimal control. When σ decreases, the variance of the uncertain asset decreases, and investing in the uncertain



(a) Value function for $t = 0$, $\gamma = 0.8$, case B

(b) Distribution of π for $0 \leq t \leq T$

Fig. 15.6: Plots for reference case A, where γ is increased from 0.3 to 0.8. Compare with Figures 15.1a and 15.3a.



(a) Value function for $t = 0$, case B

(b) Consumption

Fig. 15.7: Plots for reference case A, where D is increased from 1 to 100. Compare with Figures 15.1a and 15.4a.

asset becomes less risky. The agent will invest more in the risky asset, and the value of π increases.

Figure 15.9 shows how the value of π varies with $\hat{r}$ and $\hat{\mu}$. When $\hat{r}$ is close to $\hat{\mu}$, the expected return from the safe asset is almost as large as the expected return from the uncertain asset, and the investor will invest almost all her wealth in the safe asset. When $\hat{\mu}$ is large, on the other hand, the expected return from the uncertain asset is so large that the agent will invest much in the uncertain asset, despite the larger risk.

Figure 15.10 shows how the value of β influences the value function and the optimal control. A large value of β signalizes that the investor gets more satisfaction from recent consumption compared to earlier consumption. The satisfaction the agent gets from consumption at a certain time is "forgotten" fast. Consumption close to T will decrease, as this consumption contributes less to final satisfaction than before; it is better to save

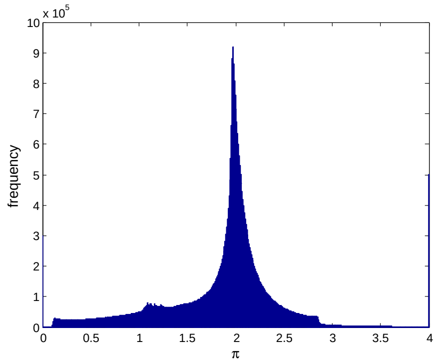
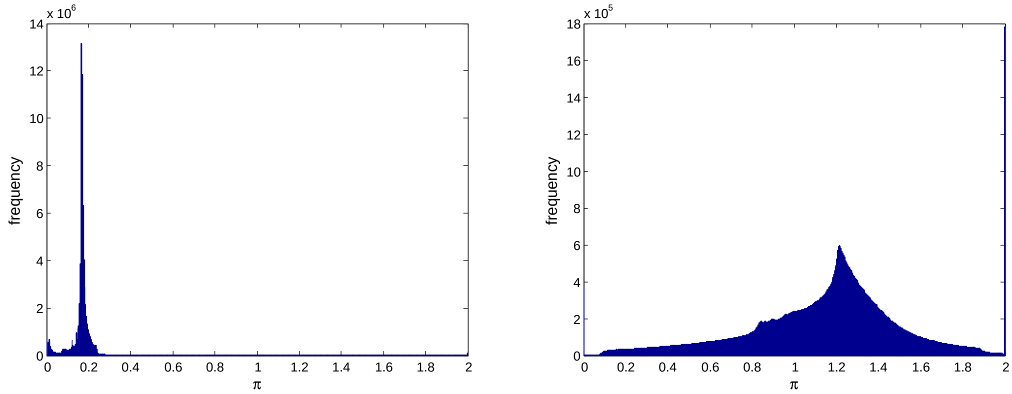


Fig. 15.8: Plot for reference case A, where σ is decreased from 0.3 to 0.15. Compare with Figure 15.3a.



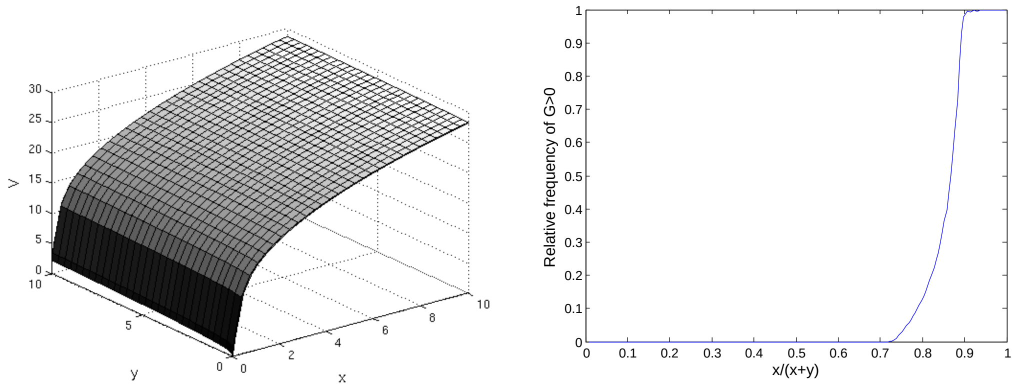
(a) $\hat{r} = 0.13$, case A

(b) $\hat{\mu} = 0.21$, case A

Fig. 15.9: Plot for reference case A, where the values of $\hat{r}$ and $\hat{\mu}$, respectively, are increased. Compare with Figure 15.3a.

wealth, and then do one large consumption gulp at time T. This change of strategy is shown well in our simulations, as the tail on Figure 15.4a, which corresponds to consumption close to time T, is smaller on Figure 15.10b. The value function is almost independent of y, as the the initial value of $Y^{\pi,C}$ is forgotten fast, and contributes little to the investor's satisfaction in most of the time interval $[0, T]$. The variant of the Merton problem described in Chapter 7.2, corresponds to the case $\beta \to \infty$, and in this case the value function is only a function of x.

Figure 15.11 shows that the value function not necessarily is decreasing in t. We have increased D from 1 to 200, and have decreased β from 2 to 0.2. When D is small, almost all satisfaction of the agent expressed in $V(t, x, y)$, is coming from consumption in the interval $[t, T)$. To obtain the largest possible total satisfaction, it is an advantage to invest over a long time interval. Therefore V is decreasing in t for $A \ll 1$. When $A \gg 1$, almost all satisfaction of the investor comes from the *final* values of $X^{\pi,C}$ and $Y^{\pi,C}$, and the length of the investment period becomes less important. Therefore the value function may be increasing in t for large values of D. It is not only D and β that



(a) $\hat{r} = 0.13$, case A

(b) $\hat{\mu} = 0.21$, case A

Fig. 15.10: Plot for reference case A, where the value of β is increased from 2 to 10. Compare with Figures 15.1a and 15.4a.

have a influence on the growth of V in t; also δ, $\hat{r}$ and $\hat{\mu}$ influence the result. Decreasing δ has the same effect as increasing D, as a small value of δ implies that the terminal values of $X^{\pi,C}$ and $Y^{\pi,C}$ become more important. If $\hat{r}$ and $\hat{\mu}$ are small, the expected growth of $X^{\pi,C}$ decreases, and this may give us a value function that is increasing in t.

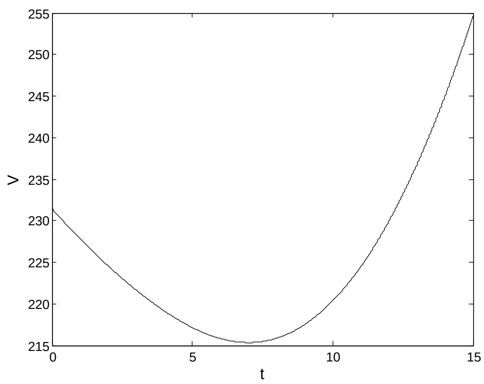
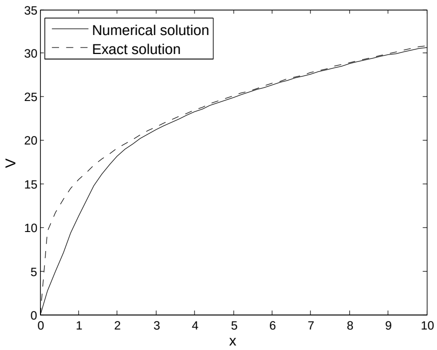


Fig. 15.11: Plot of the value function for reference case A, where D is increased from 1 to 200, and β is decreased from 2 to 0.2. We have $x, y = 5$, while t varies.

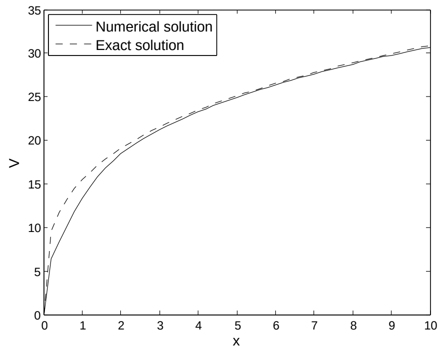
15.3 Boundary conditions

In this section we will study which boundary conditions give best simulation results for reference case B. We will also see that the boundary conditions in general have little influence on V away from the boundaries.

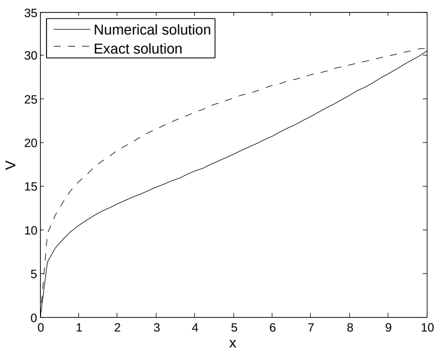
We start with the boundary condition for $y = 0$, see Figure 15.12. We see that boundary conditions (11.13) and (11.15) give best results, i.e., we should either employ the "regular" scheme on the boundary $y = 0$ or assume $G(D_X V) = 0$. Boundary condition (11.15) does not work as well as (11.13) for small x. This is natural, because (11.15) is associated with non-zero consumption, and for sufficiently small x, grid points associated with no consumption will be involved in the calculation of $\tilde{v}_{i,0}^m$. For large x, however, the two schemes (11.13) and (11.15) perform equally well. From the boundary condition $V \equiv 0$ on Figure 15.12d, we see that the boundary condition do not have much impact on V in the interior of the domain. The boundary condition (11.12) estimates V to be too small. The boundary condition (11.13) will be employed in all other simulations in this thesis.



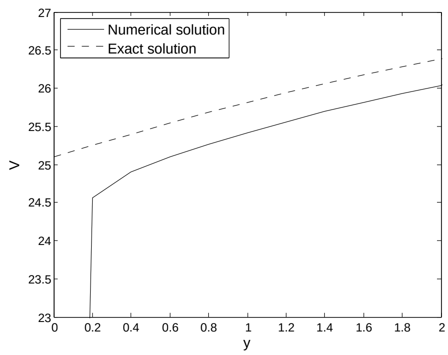
(a) Boundary condition (11.15), $y = 0, t = 0$



(b) Boundary condition (11.13), $y = 0, t = 0$



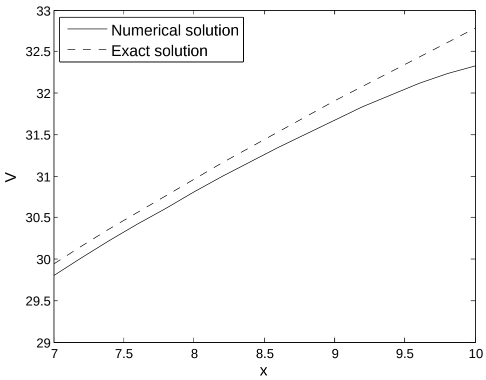
(c) Boundary condition (11.12), $y = 0, t = 0$



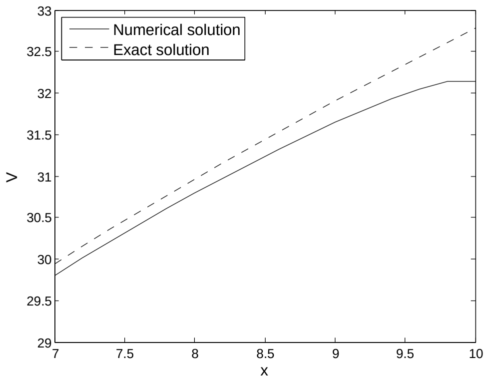
(d) Boundary condition $V \equiv 0$, $x = 5, t = 0$

Fig. 15.12: Plots of the value function for reference case B, where various boundary conditions for $y = 0$ are compared. We have $\Delta x = 0.2$ and $T = 5$ at all the plots.

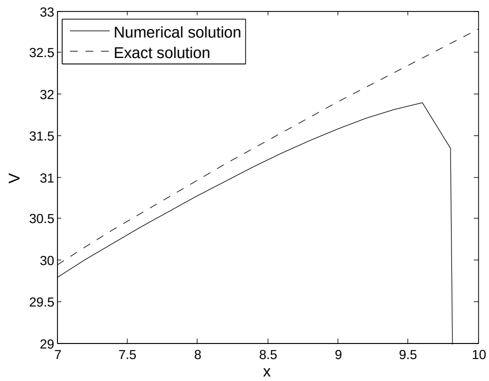
Now we consider the boundary $x = x_{max}$, see Figure 15.13. We see from Figure 15.13c that the boundary condition has little influence on the value function in the interior of the domain. We see that (11.17) is the best boundary condition, as the difference between the exact and simulated solution is smallest in this case. The boundary condition (11.17) was derived by assuming it is optimal to consume wealth when $x = x_{max}$. The Neumann boundary condition (11.16) also gives a satisfactory result.



(a) Boundary condition (11.17).



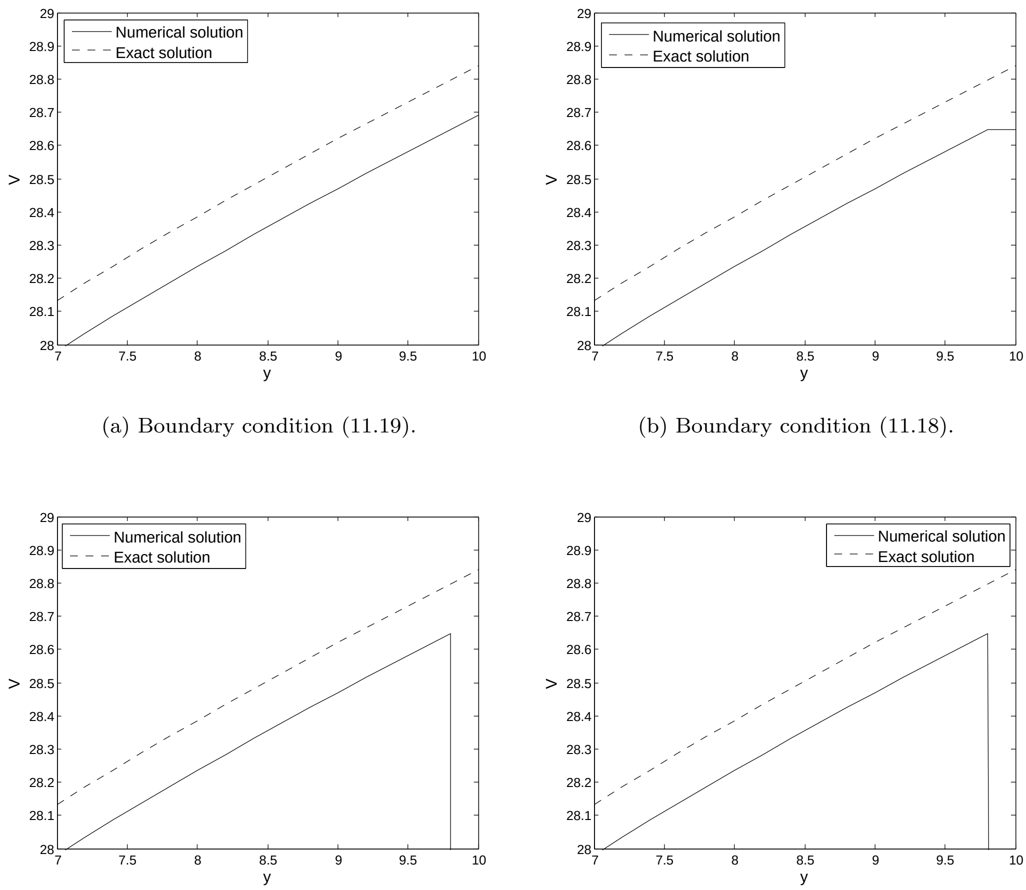
(b) Boundary condition (11.16).



(c) Boundary condition $V \equiv 0$ for $x = x_{max}$.

Fig. 15.13: Plots of the value function for reference case B, where various boundary conditions for $x = x_{max}$ are compared. We have $\Delta x = 0.2$ and $T = 5$ at all the plots. All other parameters are as given in Table 15.1. The plots show the value function as a function of x for $y = 5$ and $t = 0$.

The boundary condition that gives best simulation results for $y = y_{max}$, is (11.19), see Figure 15.14. This boundary condition is derived from the assumption that there is no consumption for $y = y_{max}$. The Neumann boundary condition (11.18) also performs relatively well. Boundary condition (11.20), which is based on the assumption that we may ignore the increase in $Y^{\pi,C}$ due to consumption, does not give a satisfactory result. It might work better for shorter time intervals, smaller values of β or smaller values of x, but for the case we are considering, $\tilde{v}$ becomes much smaller than the analytical solution.



(a) Boundary condition (11.19).

(b) Boundary condition (11.18).

(c) Boundary condition (11.20).

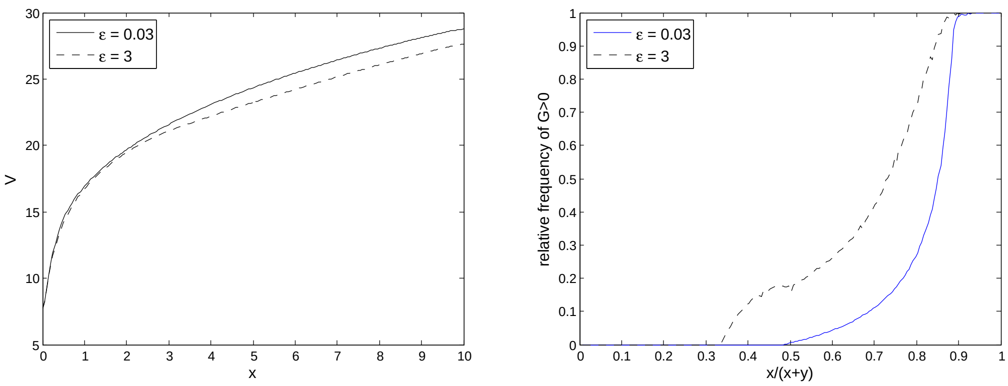
(d) Boundary condition $V \equiv 0$ for $y = y_{max}$.

Fig. 15.14: Plots of the value function for reference case B, where various boundary conditions for $x = x_{max}$ are compared. We have $\Delta x = 0.2$ and $T = 5$ at all the plots. The plots show the value function as a function of y, and $x = 5$ and $t = 0$ for all the plots.

15.4 Singular versus non-singular problem

Our MATLAB script does not solve the singular optimization problem directly, but solves the non-singular optimization problem for a user-given value of ϵ. This gives us the opportunity to compare properties of the singular problem to properties of the non-singular problems.

We assume $\epsilon = 0.03$ is a so small value of ϵ, that the produced graphs are sufficiently close to the singular problem, and we compare results for $\epsilon = 0.03$ and $\epsilon = 3$. Figure 15.15a shows that the value function of the non-singular problem, is smaller than the value function of the singular problem. This corresponds well with Theorem 9.11. We see on Figure 15.15b that there is consumption for smaller values of $x/(x + y)$ when $\epsilon = 3$. When ϵ increases, the consumption rate is generally smaller, and therefore we need to consume more frequently in order to obtain a sufficiently large total consumption.



(a) Value function for $t = 0, y = 5$ (b) Consumption as a function of $x/(x + y)$

Fig. 15.15: These plots compare the simulation results for two values of ϵ. All parameter values except for ϵ are as described in Table 15.1 for case A. The simulation results for $\epsilon = 0.03$ are almost identical to the results for the singular problem, while the results for $\epsilon = 3$ corresponds to results for a non-singular problem.

15.5 Rate of convergence and computing time

In this section we will calculate an approximate rate of convergence for our finite difference scheme, and study its running time. We will perform the analysis for reference case B, as we have an explicit solution formula for this case.

We define the error E_Δ of the scheme as

$$E_\Delta = \max_{\substack{0 \le i < 0.8N_x, \\ 0 \le j < 0.8N_y, \\ 0 \le m \le N_t}} |\tilde{v}_{i,j}^m - V(x_{i,j}^m)|,$$

where $\Delta = (\Delta x, \Delta t, \epsilon)$ and we have assumed $\Delta y = \Delta x$. Note that we do not include points close to the boundaries $x = x_{max}$ and $y = y_{max}$, because we do not want possibly inaccurate boundary conditions for large x and y to disturb the result. In case the scheme is inaccurate at these boundaries, it is easy to obtain an accurate value at a certain point by increasing the domain size slightly. However, we do include points close to the boundaries $x = 0$ and $y = 0$, as possible inaccuracies here are more serious; we cannot get rid of inaccuracies at $x = 0$ and $y = 0$ by doing small adjustments on the grid.

We assume the error can be written as

$$E_\Delta = C_{\Delta x} \Delta x^{r_{\Delta x}} + C_{\Delta t} \Delta t^{r_{\Delta t}} + C_\epsilon \epsilon^{r_\epsilon}$$

for constants $r_{\Delta x}$, $r_{\Delta t}$, r_ϵ, $C_{\Delta x}$, $C_{\Delta t}$ and C_ϵ. We wish to find approximate expressions for $r_{\Delta x}$, $r_{\Delta t}$, r_ϵ, $C_{\Delta x}$, $C_{\Delta t}$ and C_ϵ. Assume $\Delta_1 = (\Delta x_1, \Delta t, \epsilon)$, $\Delta_2 = (\Delta x_2, \Delta t, \epsilon)$ and $\Delta_3 = (\Delta x_3, \Delta t, \epsilon)$, where $\Delta x_1/\Delta x_2 = \Delta x_2/\Delta x_3 = a$ for some constant $a > 1$. Then we have

$$E_{\Delta_1} - E_{\Delta_2} = C_{\Delta x} \Delta x_1^{r_{\Delta x}} \left(1 - (\Delta x_2/\Delta x_1)^{r_{\Delta x}}\right)$$

and

$$E_{\Delta_2} - E_{\Delta_3} = C_{\Delta x} \Delta x_2^{r_{\Delta x}} \left(1 - (\Delta x_3/\Delta x_2)^{r_{\Delta x}}\right),$$

which implies that

$$\frac{E_{\Delta_1} - E_{\Delta_2}}{E_{\Delta_2} - E_{\Delta_3}} = a^{r_{\Delta x}},$$

or

$$r_{\Delta x} = \frac{\log\left(\frac{E_{\Delta_1} - E_{\Delta_2}}{E_{\Delta_2} - E_{\Delta_3}}\right)}{\log a}. \tag{15.3}$$

As soon as we have determined $r_{\Delta x}$, it is easy to find an approximate expression for $C_{\Delta x}$:

$$C_{\Delta x} = \frac{E_{\Delta_1} - E_{\Delta_2}}{\Delta x_1^{r_{\Delta x}} \left(1 - (\Delta x_2/\Delta x_1)^{r_{\Delta x}}\right)}. \tag{15.4}$$

Equations (15.3) and (15.4) will be employed when we calculate approximate expressions for $C_{\Delta x}$ and $r_{\Delta x}$. We get similar equations for $C_{\Delta t}$, C_ϵ, $r_{\Delta t}$ and r_ϵ. For Δx and Δt we will let $a = 2$, while we will let $a = 3$ for ϵ. Tables 15.2-15.4 show our calculated values for $C_{\Delta x}$, $C_{\Delta t}$, C_ϵ, $r_{\Delta x}$, $r_{\Delta t}$ and r_ϵ. We see that an approximate expression for the error E_Δ, is

$$E_\Delta = 0.667\Delta x^{0.300} + 0.0127\Delta t + 0.00166\epsilon,$$

and that the error is dominated by the Δx term.

Δx	E_Δ	ΔE_Δ	$r_{\Delta x}$	$C_{\Delta x}$	Running time (s)
0.400	1.0981				23.9
0.200	0.8918	0.2063		0.6670	49.4
0.100	0.7243	0.1675	0.3006	0.6668	116.6
0.050	0.5883	0.1360	0.3003	0.6666	450.3
0.025	0.4778	0.1105	0.3002	0.6665	3443.3

Table 15.2: This table gives an approximated rate of convergence in Δx and Δy for reference case B. The running time of the program is also given. We had $\Delta y = \Delta x$ in all the simulations. The value of Δx and Δy varies, while $\Delta t = 5 \cdot 10^{-5}$ and $\epsilon = 0.03$ for all the runs. We have $\Delta E_\Delta = E_{\Delta_1} - E_{\Delta_2}$, where Δx_2 is the value of Δx in the current row, while Δx_1 is the value of Δx for the row above. The value of $r_{\Delta x}$ is calculated by using (15.3), and the value in a specific row is calculated on basis of the error in the current row, in addition to the error in the two rows above. The value of $C_{\Delta x}$ is calculated by using (15.4) with $r_{\Delta x} = 0.3$. All parameter values are as given in Table 15.1, expect that we have $T = 1$.

We can justify this expression from a theoretical point of view by studying the truncation error (12.4). We assume the viscosity solution V_ϵ of (11.1) is a classical solution, and that ϵ is so small that V_ϵ and all its derivatives may be approximated by V and the corresponding derivatives of V, where V is given by (7.27). For smooth solutions of (11.1), we can obtain an approximate rate of convergence by inserting the exact solution into the expression for the truncation error, see the discussion after Lemma 12.6. The truncation error associated with Δt in (12.5), is of order

$$O\big(|V_{tt}|_{L^\infty(\mathcal{D}_T)}\Delta t\big) = O(\Delta t),$$

because V_{tt} is bounded in our computational domain $\mathcal{D}_T$, see (7.27). This corresponds well with a rate of convergence in Δt equal to 1.

Δt	E_Δ	ΔE_Δ	$r_{\Delta t}$	$C_{\Delta t}$	Running time (s)
$5.000 \cdot 10^{-4}$	0.724436				15.1
$2.500 \cdot 10^{-4}$	0.724432	$3.1705 \cdot 10^{-6}$		0.012682	26.6
$1.250 \cdot 10^{-4}$	0.724431	$1.5858 \cdot 10^{-6}$	0.9995	0.012686	51.8
$6.250 \cdot 10^{-5}$	0.724430	$7.9295 \cdot 10^{-7}$	0.9999	0.012687	100.5
$3.125 \cdot 10^{-5}$	0.724430	$3.9646 \cdot 10^{-7}$	1.0001	0.012687	178.9

Table 15.3: This table displays the same information as Table 15.2, except that we have varied Δt instead of varying Δx and Δy. The value of both Δx, Δy and ϵ were 0.1 in all the runs. We assumed $r_{\Delta t} = 1$ when calculating $C_{\Delta t}$.

ϵ	E_Δ	ΔE_Δ	r_ϵ	C_ϵ	Running time (s)
0.81	0.72553				30.0
0.27	0.72471	$8.220 \cdot 10^{-4}$		0.001522	29.5
0.09	0.72442	$2.904 \cdot 10^{-4}$	0.9471	0.001613	31.1
0.03	0.72432	$9.868 \cdot 10^{-5}$	0.9824	0.001645	30.0
0.01	0.72428	$3.316 \cdot 10^{-5}$	0.9941	0.001655	29.9

Table 15.4: This table displays the same information as Table 15.2, except that we have varied the value of ϵ instead of varying Δx. The values of Δx, Δy and Δt were 0.1, 0.1 and $2 \cdot 10^{-4}$, respectively, in all the runs. We assumed $r_\epsilon = 1$ when calculating C_ϵ.

The truncation error associated with Δx in (12.5) is of order

$$O\left(\Delta x |V_{xx}|_{L^\infty(\overline{\mathcal{D}_T})} + \Delta x |V_{yy}|_{L^\infty(\overline{\mathcal{D}_T})}\right)$$

for $\Delta x = \Delta y$ and $\nu \equiv 0$. However, since the second derivatives of the exact solution V are unbounded, the coefficients $|V_{xx}|_{L^\infty(\overline{\mathcal{D}_T})}$ and $|V_{yy}|_{L^\infty(\overline{\mathcal{D}_T})}$ are non-existent. We assume that we may replace $|V_{xx}|_{L^\infty(\mathcal{D}_T)}$ and $|V_{yy}|_{L^\infty(\mathcal{D}_T)}$ by $|V_{xx}|_{L^\infty(\mathcal{H}_T^*)}$ and $|V_{yy}|_{L^\infty(\mathcal{H}_T^*)}$, respectively, i.e., we only consider values of V_{xx} and V_{yy} at grid points away from the boundary $x = 0$. Doing this assumption, we get a truncation error in Δx of the following order:

$$O\left(\Delta x |V_{xx}|_{L^\infty(\mathcal{H}_T^*)} + \Delta x |V_{yy}|_{L^\infty(\mathcal{H}_T^*)}\right). \tag{15.5}$$

This expression is *not* bounded by $K\Delta x$, because V_{xx} and V_{yy} are not bounded on $\mathcal{H}_T^*$ when $\Delta x \to 0$. However, we manage to justify that the rate of convergence is 0.3 by deriving a new expression for the truncation error, and by considering the order of which the second derivatives of the exact solution converges to infinity. We want to prove that the rate of convergence is γ, which will imply that the rate of convergence is 0.3 in our case, see Table 15.1. The second derivatives of the exact solution are given by

$$V_{xx}(t,x,y) = \begin{cases} -e^{-\delta t} k_2 k^{-\rho} \rho(1-\rho) y^{\gamma-\rho} x^{\rho-2} & \text{for } x \le ky, \\ -\frac{k_3 \gamma(1-\gamma)\beta^2}{(1+\beta k)^2} e^{-\delta t} \left(\frac{y+\beta x}{1+\beta k}\right)^{\gamma-2} & \text{for } x > ky, \end{cases}$$

and

$$V_{yy}(t,x,y) = \begin{cases} -e^{-\delta t}\Big(k_1 \gamma(1-\gamma) y^{\gamma-2} \\ \quad + k_2 k^{-\rho}(\gamma-\rho)(1+\rho-\gamma) x^\rho y^{\gamma-\rho-2}\Big) & \text{for } x \le ky, \\ -\frac{k_3 \gamma(1-\gamma)}{(1+\beta k)^2} e^{-\delta t} \left(\frac{y+\beta x}{1+\beta k}\right)^{\gamma-2} & \text{for } x > ky. \end{cases}$$

We see that the magnitude of second derivatives are decreasing in x and y, so, for a given grid, the magnitude of the second derivative is biggest in one of the points $(\Delta x, l\Delta x)$ and $(\Delta x, 0)$, where l is the smallest natural number such that $\Delta x \leq kl\Delta x$. The second derivatives of V are of order $O(\Delta x^{\gamma-2})$ in $(\Delta x, l\Delta x)$ and $(\Delta x, 0)$. If we insert the expressions for V_{xx} and V_{yy} into (15.5), and evaluate the derivatives at $(\Delta x, l\Delta x)$ and $(\Delta x, 0)$, we get a truncation error of order $O(\Delta x^{\gamma-1})$, i.e., a truncation error that diverges as $\Delta x \to 0$.

We need to derive a new and better expression for the truncation error in order to get a truncation error of order Δx^{γ}. We start with the case $x \leq ky$. If we look more closely on the HJB equation (11.1) with $c = 0$, we see that there is a factor x in front of v_x, a factor x^2 in front of v_{xx}, and a factor y in front of v_y. We put $c = 0$, as we expect to have $c = 0$ for $x \leq ky$. Including the factors x and y in the expression for the truncation error, we obtain a truncation error of the following order:

$$O\Big(x\Delta x|\partial_x^2 V|_{L^\infty(\mathcal{H}_T^*)} + x^2\Delta x^2|\partial_x^4 V|_{L^\infty(\mathcal{H}_T^*)} + y\Delta x|\partial_y^2 V|_{L^\infty(\mathcal{H}_T^*)}\Big).$$

Note that the term $\Delta x^2|\partial_x^4 V|_{L^\infty(\mathcal{H}_T^*)}$, which is associated with the term $\frac{1}{2}(\sigma\pi x)^2 V_{xx}$ and was not included in (12.5), has been included now, as we know that $|\partial_x^4 V|_{L^\infty(\mathcal{H}_T^*)}$ converges faster to infinity than $|\partial_x^2 V|_{L^\infty(\mathcal{H}_T^*)}$. Evaluating this new truncation error at $(x, y) = (\Delta x, l\Delta x)$, we obtain a truncation error of order $O(\Delta x^{\gamma})$ as wanted.

Now we consider the case $x \geq ky$. We want to show that the truncation error associated with the term $\frac{1}{\epsilon}\max\big\{G(D_X V); 0\big\}$ is of order $O(\Delta x^{\gamma})$. We know that the truncation error is of order $O(\Delta x)$ when x or y are bounded away from 0, so we will focus on calculating the truncation error in the point $(t, x, y) = (t_n, \Delta x, 0)$ for some $n \leq N_t$. We know that $(\Delta x, 0)$ converges towards the origin when $|\mathcal{H}_T| \to 0$, and that $(\Delta x, 0)$ is the point of the grid $\mathcal{H}^* \cup \{x \geq ky\}$ where the second derivatives of V take their largest value. We cannot put $c = 0$ when we calculate the truncation error, as we expect to have $c = 1/\epsilon$ for $x \geq ky$. There is no factor x in front of V_x in G, and no factor y in front of V_y in G, so the approach in the paragraph above does not work. However, we see by Taylor expansions that the truncation error associated with $G(D_X V)$ always is negative, since $V_{xx}, V_{yy} \leq 0$. Therefore the term $\frac{1}{\epsilon}\max\big\{G(D_X V); 0\big\}$ is always underestimated by our numerical scheme, and the truncation error associated with the term cannot be larger than $G(D_X V)$. We know that

$$\begin{aligned} G\big(V_\epsilon(t_n, \Delta x, 0)\big) &= \epsilon\Big((V_\epsilon)_t(t_n, \Delta x, 0) + F\big(t_n, \Delta x, 0, (V_\epsilon)_x, (V_\epsilon)_{xx}, \mathscr{I}^\pi(t_n, \Delta x, 0, V_\epsilon)\big)\Big) \\ &= O(\Delta x^{\gamma}), \end{aligned}$$

so we obtain a truncation error of the wanted order.

One may argue that the real convergence rate of the scheme in x and y, is better than γ. We obtain a convergence rate of γ instead of 1, because we use grid points closer and closer to 0. When we use a finer grid, the coefficient of Δx in (15.5) is increased, because $|V_{xx}|_{L^\infty(\mathcal{H}_T^*)}$ and $|V_{yy}|_{L^\infty(\mathcal{H}_T^*)}$ are evaluated at points closer to $(0, 0)$. A more fair comparison might be to calculate the error *on the same grid* for all values of Δx. If we calculate the error on the same grid for all values Δx, we get a better impression of how the error changes at a fixed point of the domain. When we reduce the value of Δx from $2\Delta x_1$ to Δx_1, the error of the scheme is reduced by a factor of

0.5^γ, provided we have found each error by considering the grid used in the current simulation. However, the error of the scheme in point $(2\Delta x_1, 2\Delta x_1)$ may be reduced by a larger factor than 0.5^γ. Using the same 25×25 grid in all error calculations, we obtain errors of 1.0981, 0.8918, 0.4805, 0.1033 and 0.0355, respectively, for the five values of Δx shown in Table 15.2. This corresponds to a better rate of convergence than 0.3. Another approach to capture the scheme's performance at a fixed interior point of the domain, is to only consider the error on a subset of $\mathcal{D}_T$. If we define E_Δ to be the largest error for $(x_i, y_j) \in [0.1x_{max}, 0.8x_{max}] \times [0.1y_{max}, 0.8y_{max}]$, we obtain errors 0.6751, 0.2662, 0.0756, 0.0350 and 0.0171, respectively, for the five values of Δr displayed in Table 15.2. This corresponds to a rate of convergence approximately equal to 1.

Tables 15.2-15.4 also show the running time in the various runs. We see that the running time is approximately independent of ϵ, while it is linear in N_t. This corresponds well with our discussion in Chapter 14. There is no simple relation between N_x and the running time in Table 15.2. The matrix size is quadrupled from one row to the next, so we would expect the running time to be quadrupled as well from a theoretical point of view. However, as Table 15.5 shows, the time complexity of matrix operations in MATLAB are not proportional with the matrix size, not even for operations like addition and component-wise multiplication. The time requirements do a large jump for matrix sizes around 400×400, and this explains why the running time suddenly increased from 430.3 to 3443.3 when Δx was reduced from 0.05 to 0.025. The running time in Table 15.2 grows sublinearly with the matrix size for small matrices, and this can also be explained by Table 15.5.

m	Running time (s)	Ratio
25	0.0025	
50	0.0041	1.6401
100	0.0138	3.3325
200	0.0562	4.0721
400	0.6636	11.8025
800	3.1421	4.7349
1600	12.5266	3.9867
3200	42.7830	3.4154
6400	163.5863	3.8236

Table 15.5: The MATLAB code $C = A+B; D = C+A; E = D+B; F = E+A; G = F+B; H = F+A;$ has been run 60 times with different random $m \times m$ matrices, and the sum of the running times for each matrix dimension has been added. In the row called "Ratio", the ratio between the running time of the current column and the neighbouring column has been calculated. From a theoretical point of view, one would expect the ratio to always be ≈ 4, as the number of elements of a matrix is proportional to the square of its side length. We see that this is not the case, and the deviation can be explained by the time requirements associated with memory allocations in MATLAB. We see that the running time grows sublinearly with the matrix size for small matrices, but that it does a large jump for matrix sizes around 400×400. The running times vary little (< 0.1s) between repeated experiments, so the differences cannot be explained by the randomness of the method.

15.6 One-dimensional scheme: Jumps in the Lévy process

In this and the next section we will consider problems where $\nu \not\equiv 0$, and the scheme described in Chapter 13 will be employed in the simulations. In this section we will compare simulation results for problems with and without jumps. In the next section we will study the running time and the rate of convergence of the scheme. Recall that the value function for the one-dimensional problem is called $\overline{V}$, and that it satisfies $\overline{V}(t,r) = V(t,r,1)$ for all $r > 0$ and $t \in [0,T]$, see Section 5.1.

We will perform the simulations for two variants of reference cases A and B, called A' and B', respectively. Simulation results for both cases will be studied in this section, while simulation results for case B' will be studied in the next section. Case A' is identical to case A, expect that the Lévy measure has density function

$$k(z) = \lambda \frac{e^{-b|z|}}{|z|^2} \tag{15.6}$$

for constants $\lambda, b \in [0,\infty)$, instead of being identically equal to 0. Case B' is identical to case B, expect that the constants involved are different, and that the Lévy measure has density function given by

$$k(z) = \lambda \delta_{-\xi}(z) \tag{15.7}$$

for $\lambda \in [0,\infty)$ and $\xi \in (0,1)$. See table 15.6.

Case	T	δ	γ	D	σ	$\hat{r}$	$\hat{\mu}$	β	U	W	ν	λ	b	ξ
A'	15	0.10000	0.3	1	0.3	0.07	0.10	2	(15.1)	(15.2)	(15.6)	100	0.1	-
B'	1	0.18751	0.4	-	0.2	0.04	0.14	1	(15.1)	(7.33)	(15.7)	10	-	0.3

Table 15.6: This table describes the cases A' and B'. The columns U, W and ν refer to equations.

We will use case A' to study what impact the jump has on the value function and the choice of optimal controls, and we will compare the value functions calculated with the two schemes (11.21) and (13.1). The specific choice of ν was chosen to demonstrate that the numerical scheme is stable even for unbounded Lévy measures. As described in [12], unbounded Lévy measures are more difficult to handle numerically than finite Lévy measures. The choice of ν is also realistic from a financial point of view, as the probability of large jumps is smaller than the probability of small jumps. We will use case B' to study the rate of convergence of the scheme, see Section 15.7, and to check that our program reproduces the correct value function, values of π and consumption strategy.

First we check that the one-dimensional scheme gives approximately the same simulation results as the two-dimensional scheme. We see on Figure 15.16 that the calculated value function is approximately similar for the two schemes. We also note that the value of π on Figure 15.17c is approximately equal to the value of π on Figure 15.3a, and that the consumption strategy on Figure 15.17b is approximately equal to the consumption strategy on Figure 15.4a. We can conclude that the two schemes seem to give the same simulation results.

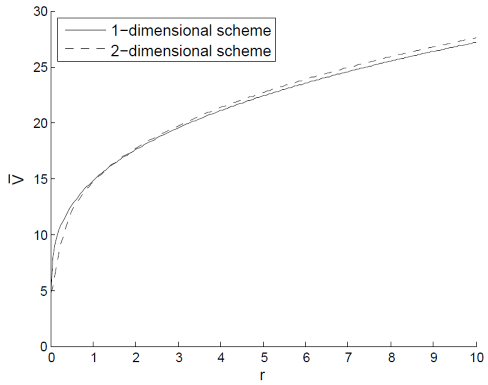


Fig. 15.16: The value function $\overline{V}$ for case A calculated with the two schemes (11.21) and (13.2) for $t = 0$. We have used $\Delta r = 0.05$ in the one-dimensional scheme, otherwise all constants are as specified in Table 15.1.

Now we compare the simulation results for cases A and A', see Figure 15.17. The one-dimensional scheme described in Chapter 13 is employed in both simulations. The only difference between the problems, is that we have added a jump process to the Lévy process in A', while the Lévy measure in case A is $\nu \equiv 0$. The expected return of the stock is the same for cases A and A', but the uncertainty in case A' is larger. The consequence of the increase in uncertainty, is that the agent invests all her wealth in the safe asset, see Figure 15.17d. We also see that the value function is slightly lower in case A', which is natural, as the investor decreases the expected return of her investments in order to reduce volatility. The consumption strategy is equal in the two cases.

See Figure 15.18 for simulation results for case B'. The parameters in Table 15.6 imply that $\pi = 0.3$ and $k/(k+1) \approx 0.7388$. For comparison, we would have had $\pi = 4.6$ and $k/(k+1) \approx 0.8970$ in case B' if $\nu \equiv 0$. We see that our program gives us the expected value of π and the expected consumption strategy, and that the addition of a jump clearly has had an effect on the choice of optimal controls.

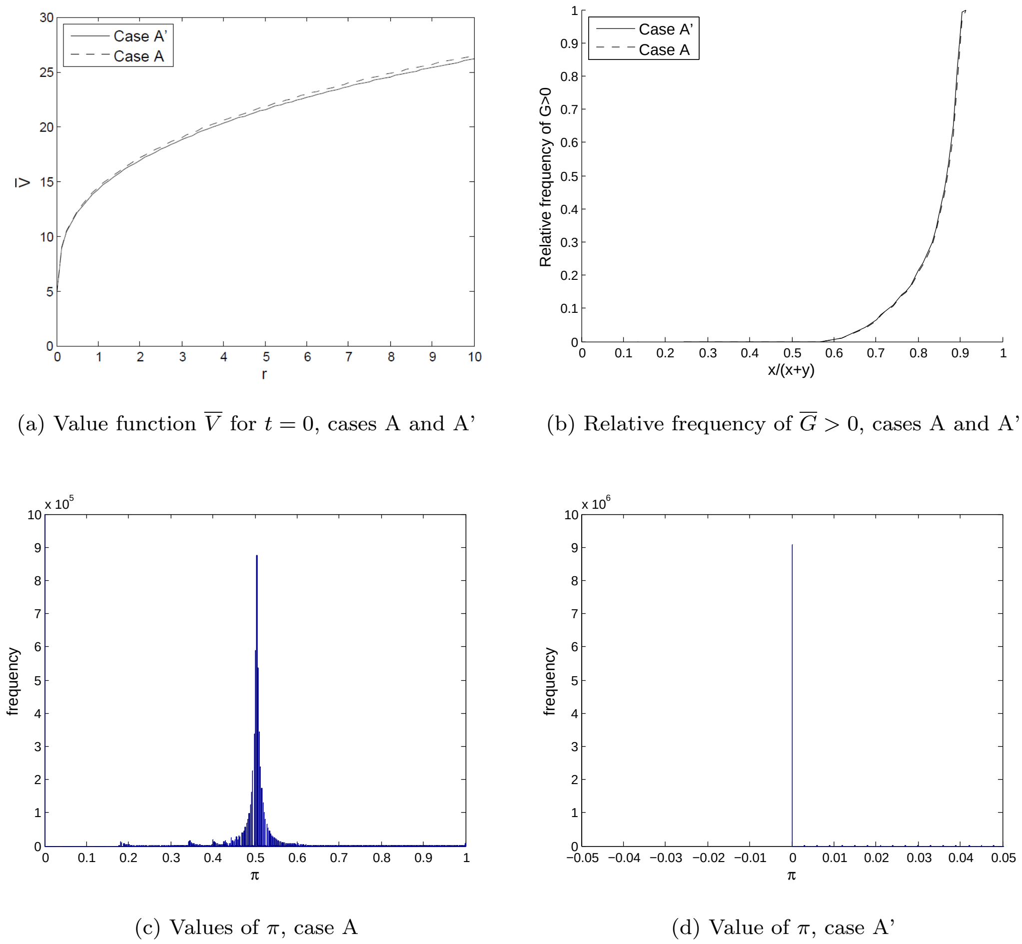
15.7 One-dimensional scheme: Computing time and rate of convergence

In this section we will calculate an approximate rate of convergence for the scheme (13.2), and we will study its running time. The analysis is performed for case B', as we have an explicit solution formula for this case.

We define the error E_Δ of the scheme as

$$E_\Delta = \max_{\substack{0 \le i < 0.5N_r, \\ 0 \le m \le N_t}} |\tilde{v}_i^m - \overline{V}(r_i^m)|,$$

where $\Delta = (\Delta r, \Delta t, \epsilon)$. Note that we only include points corresponding to $r \le r_{max}/2$. The reason for this, is that the truncation of the sum (11.10) makes the scheme inaccurate for large values of r.



(a) Value function $\overline{V}$ for $t = 0$, cases A and A'

(b) Relative frequency of $\overline{G} > 0$, cases A and A'

(c) Values of π, case A

(d) Value of π, case A'

Fig. 15.17: Simulation results for cases A and A', using the one-dimensional scheme described in Chapter 13. Both simulations are performed with $\Delta t = 1.3 \cdot 10^{-4}$, $\Delta r = 0.125$ and $\epsilon = 0.03$, otherwise all constants and functions are as described in Tables 15.1 and 15.6.

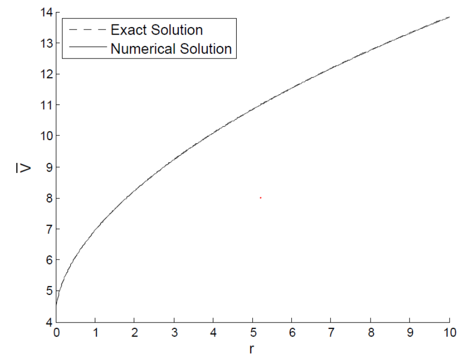
Just as for the two-dimensional scheme, the rate of convergence in Δt and ϵ is 1. See Tables 15.10 and 15.11. When there are no jumps in the Lévy process, the rate of convergence in Δr is approximately equal to $\rho \approx 0.3369$, see Table 15.9. Just as for the singular case, the rate of convergence in Δt and Δr can be explained by calculating the truncation error of the scheme. Assuming we only need to consider values of $\overline{V}$ and its derivatives at $\mathcal{H}'_T$, we see that the truncation error is of order

$$O\Big(\big|\partial_t^2 \overline{V}\big|_{L^\infty(\mathcal{H}'_T)} \Delta t + r\big|\partial_r^2 \overline{V}\big|_{L^\infty(\mathcal{H}'_T)} \Delta r + r^2\big|\partial_r^4 \overline{V}\big|_{L^\infty(\mathcal{H}'_T)} \Delta r^2 + c\big|\partial_r^2 \overline{V}\big|_{L^\infty(\mathcal{H}'_T)} \Delta r\Big).$$

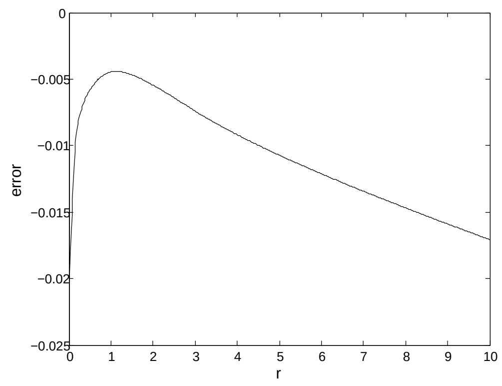
For $r \in [k, r_{max}]$, all derivatives of $\overline{V}$ are bounded by (7.27), so the truncation error is of order $O(\Delta t + \Delta r)$. For $r \in [0, k]$, the derivatives of $\overline{V}$ with respect to r are unbounded:

$$\begin{aligned}
\partial_r^2 \overline{V}(t, r) =& -e^{-\delta t} k_2 k^{-\rho} \rho(1-\rho) r^{\rho-2} = O(r^{\rho-2}),\\
\partial_r^4 \overline{V}(t, r) =& -e^{-\delta t} k_2 k^{-\rho} \rho(1-\rho)(2-\rho)(3-\rho) r^{\rho-4} = O(r^{\rho-4}).
\end{aligned}$$

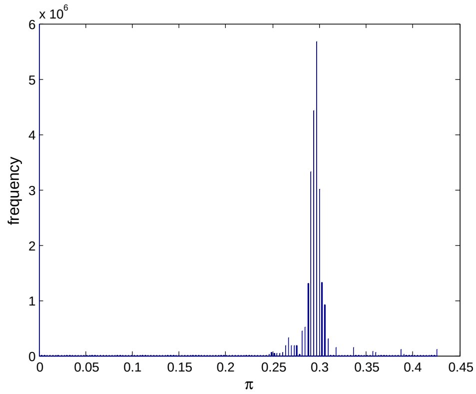
We can assume $c = 0$ for $r \leq k$, since it is not optimal to consume for $x \leq ky$, and we obtain a truncation error of order



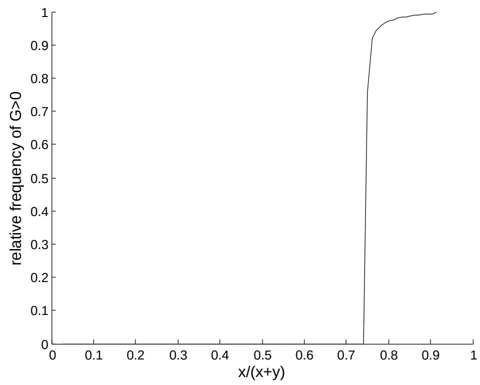
(a) The value function for $t = 0$



(b) Difference between numerical solution and exact solution



(c) Values of π, case B'



(d) Relative frequency of $G > 0$

Fig. 15.18: Simulation results for case B'. We have $\Delta t = 7 \cdot 10^{-6}$, $\Delta r = 0.01768$ and $\epsilon = 0.01$ for all the plots.

$$O\Big(\Delta t + \Delta r \cdot \Delta r^{\rho-2} \cdot \Delta r + \Delta r^2 \cdot \Delta r^{\rho-4} \cdot \Delta r^2\Big) = O(\Delta t + \Delta r^{\rho}).$$

This estimate fits the numerical experiments in Tables 15.9 and 15.10 well.

In Table 15.8 we have tried to calculate a rate of convergence in Δr for a problem with jumps in the Lévy process. For $\Delta r \geq 0.1$, the largest error occurs for $r = r_{max}/2$, but for $\Delta r \leq 0.07071$, the largest error occurs for $r = \Delta r$. For this reason, we do not obtain the same value of $r_{\Delta r}$ for all values of Δr. For large r, the error is dominated by the fact that we have to truncate the domain in order to perform the simulations, and for small r the error comes from the truncation error of the scheme. See Figure 15.18b, where the error is plotted as a function of r. The value of ρ associated with case B' is $\rho \approx 0.5567$, and we see that this is approximately equal to the rate of convergence we obtained for small values of Δr. For comparison, the value of ρ for case B' with $\nu \equiv 0$, is $\rho \approx 0.4618$, so the inclusion of a jump has clearly had an effect on the rate of convergence.

In Table 15.9, we have tried to calculate a rate of convergence for points in the interior of the domain $[0, r_{max}]$. We have focused on values of r in $[0.1r_{max}, 0.5r_{max}]$, to avoid errors associated with the truncation of the domain, and errors associated with using gradually smaller values of r when calculating the error. We see that the rate of convergence in Δr is approximately equal to 1. The truncation error for any $r \in [0.1r_{max}, 0.5r_{max}]$ and $t \in [0, T]$, is of order

$$O\Big(\partial_t^2 \overline{V}(t,r)\Delta t + \partial_r^2 \overline{V}(t,r)\Delta r + \partial_r^4 \overline{V}(t,r)\Delta r^2\Big).$$

By using that $\partial_r^2 \overline{V}$ and $\partial_r^4 \overline{V}$ are bounded for $r \in [0.1r_{max}, 0.5r_{max}]$, we see that the expected rate of convergence in $[0.1r_{max}, 0.5r_{max}]$ is 1 in both Δr and Δt.

We see that the running time is dominated by the calculation of the integral term in the HJB equation. Comparing Tables 15.7 and 15.9, we see that introducing a jump makes the running time 70 times larger for $\Delta r = 0.01767$, and 7.4 times larger for $\Delta r = 0.2857$. For the case $\nu \equiv 0$, there is no simple relation between Δr and the running time, by the same reason as for the two-dimensional scheme. For the case $\nu \not\equiv 0$, however, the running time is approximately linear in N_r for large N_r. The calculation of the integral term of the HJB equation employs some element-wise operations (i.e., not matrix operations), and these seem to dominate the running time for large values of N_r. Therefore we see a linear relation between N_r and the running time. There is also a linear relation between N_t and the running time, and the running time is independent of ϵ. These results are consistent with our discussion in Chapter 14.

Δr	E_Δ	ΔE_Δ	$r_{\Delta r}$	$C_{\Delta r}$	Running time (s)
0.2828	0.1402				406.3
0.2000	0.1041	$3.612 \cdot 10^{-2}$		2.8158	448.7
0.1414	0.0753	$2.875 \cdot 10^{-2}$	0.6585	2.7118	544.9
0.1000	0.0559	$1.941 \cdot 10^{-2}$	1.1344	2.2146	628.0
0.0707	0.0437	$1.227 \cdot 10^{-2}$	1.3224	1.6945	1112.6
0.0500	0.0360	$7.706 \cdot 10^{-3}$	1.3423	1.2876	1661.0
0.0354	0.0297	$6.307 \cdot 10^{-3}$	0.5780	1.2752	2129.8
0.0250	0.0245	$5.198 \cdot 10^{-3}$	0.5581	1.2716	4304.5
0.0177	0.0202	$4.294 \cdot 10^{-3}$	0.5514	1.2710	8009.0

Table 15.7: This table contains the error and running time for different values of Δr for case B'. The values of Δt and ϵ are $7 \cdot 10^{-6}$ and 0.01, respectively, for all the runs. The values of E_Δ, ΔE_Δ and $r_{\Delta r}$ are calculated as for Table 15.2, except that we only have considered values of $r < R_{max}/2$ when calculating E_Δ. All parameter values are as given in Table 15.6, expect that we have $T = 1$. We assumed $r_{\Delta r} = 0.55$ when calculating $C_{\Delta r}$.

Δr	E_Δ	ΔE_Δ	$r_{\Delta x}$	$C_{\Delta r}$	Running time (s)
0.2828	0.14021				406.3
0.2000	0.10409	$3.612 \cdot 10^{-2}$		8.30	448.7
0.1414	0.07534	$2.875 \cdot 10^{-2}$	0.6585	9.34	544.9
0.1000	0.05593	$1.941 \cdot 10^{-2}$	1.1344	8.92	628.0
0.0707	0.04013	$1.580 \cdot 10^{-2}$	0.5923	10.27	1112.6
0.0500	0.02907	$1.106 \cdot 10^{-2}$	1.0297	10.17	1661.0
0.0354	0.02090	$8.165 \cdot 10^{-3}$	0.8757	10.61	2129.8
0.0250	0.01499	$5.912 \cdot 10^{-3}$	0.9319	10.87	4304.5
0.0177	0.01072	$4.278 \cdot 10^{-3}$	0.9335	11.12	8009.0

Table 15.8: This table displays the exact same information as table 15.7, except that we only have used values of $r \in [0.1R_{max},\, 0.5R_{max}]$ when calculating E_Δ.

Δr	E_Δ	ΔE_Δ	$r_{\Delta r}$	$C_{\Delta r}$	Running time (s)
0.2857	0.5875				54.8
0.2000	0.5209	0.06662		1.7743	55.7
0.1408	0.4628	0.05805	0.3972	1.7392	58.0
0.1000	0.4124	0.05044	0.4054	1.6966	61.3
0.0709	0.3673	0.04508	0.3243	1.6977	68.3
0.0500	0.3265	0.04081	0.2872	1.7214	76.8
0.0353	0.2905	0.03604	0.3588	1.7060	83.6
0.0250	0.2585	0.03195	0.3462	1.6968	95.0
0.0177	0.2300	0.02852	0.3277	1.6978	114.9
0.0125	0.2047	0.02532	0.3462	1.6886	141.3

Table 15.9: This table displays the exact same information as Table 15.7, except that we have performed simulations with case B instead of case B'. The one-dimensional scheme from Chapter 13 is used.

Δt	E_Δ	ΔE_Δ	$r_{\Delta t}$	$C_{\Delta t}$	Running time (s)
$1.00 \cdot 10^{-4}$	0.075352				38.2
$5.00 \cdot 10^{-5}$	0.075344	$7.631 \cdot 10^{-6}$		0.21789	73.3
$2.50 \cdot 10^{-5}$	0.075340	$3.815 \cdot 10^{-6}$	0.99997	0.21789	144.7
$1.25 \cdot 10^{-5}$	0.075338	$1.908 \cdot 10^{-6}$	0.99999	0.21790	291.4
$6.25 \cdot 10^{-6}$	0.075337	$9.539 \cdot 10^{-7}$	0.99999	0.21790	585.2

Table 15.10: This table displays the exact same information as Table 15.7, except that we have varied Δt, and let Δr and ϵ be constant. The values of Δr and ϵ are 0.1408 and 0.01, respectively. We have assumed $r_{\Delta t} = 1$ when calculating $C_{\Delta t}$.

ϵ	E_Δ	ΔE_Δ	r_ϵ	C_ϵ	Running time (s)
0.640	0.08641				66.0
0.320	0.07838	$8.032 \cdot 10^{-3}$		0.04778	65.4
0.160	0.07651	$1.869 \cdot 10^{-3}$	2.1038	0.02223	66.9
0.080	0.07584	$6.721 \cdot 10^{-4}$	1.4751	0.01599	66.5
0.040	0.07555	$1.349 \cdot 10^{-4}$	1.2144	0.01379	66.7
0.020	0.07541	$2.568 \cdot 10^{-4}$	1.1023	0.01284	67,2
0.010	0.07535	$6.516 \cdot 10^{-5}$	1.0500	0.01240	66.3
0.005	0.07531	$3.203 \cdot 10^{-5}$	1.0247	0.01219	65.3

Table 15.11: This table displays the exact same information as Table 15.7, except that we have varied ϵ instead of Δr. We used $\Delta r = 0.1408$ and $\Delta t = 10^{-4}$ in all the simulations. We have assumed $r_\epsilon = 1$ when calculating C_ϵ.

References

1. R. Almgren and A. Tourin. Optimal soaring with Hamilton-Jacobi-Bellman equations. August 2004.
2. O. Alvarez. A singular stochastic control problem in an unbounded domain. *Communications in Partial Differential Equations*, 19(11):2075–2089, 1994.
3. T.M. Apostol. *Mathematical Analysis*. Addison Wesley, second edition, 1974.
4. G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance theory. In L.C.G Rogers and D. Talay, editors, *Numerical methods in finance*, pages 1–21. Cambridge University Press, 1997.
5. G. Barles and P.E. Souganidis. Convergence of approximation schemes for parabolic Hamilton-Jacobi-Bellman equations. *Asymptotic Anal.*, 4(3):271–283, 1991.
6. O.E. Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and Stochastics*, 2:41–68, 1998.
7. F.E. Benth, K.H. Karlsen, and K. Reikvam. On the existence of optimal controls for a singular stochastic control problem in finance. In *Mathematical finance (Konstanz, 2000)*, Trends Math., pages 79–88. Birkhäuser Verlag, Basel, 2001.
8. F.E. Benth, K.H. Karlsen, and K. Reikvam. Optimal porfolio management rules in a non-Gaussian market with durability and intertemporal substitution. *Finance Stoch.*, 5(4):447–467, 2001.
9. F.E. Benth, K.H. Karlsen, and K. Reikvam. Optimal portfolio selection with consumption and nonlinear integro-differential equations with gradient constraint: a viscosity solution approach. *Finance and Stochastics*, 5:275–303, 2001.
10. F.E. Benth, K.H. Karlsen, and K. Reikvam. A semilinear Black and Scholes partial differential equation for valuing American options: approximate solutions and convergence. *Interfaces and Free Boundaries*, 6(4):379–404, 2004.
11. P. Billingsley. *Probability and Measure*. John Wiley and Sons, third edition, 1995.
12. I.H. Biswas, E.R. Jakobsen, and K.H. Karlsen. Difference-quadrature schemes for nonlinear degenerate parabolic integro-PDE. *SIAM J. Numer. Anal.*, 48(3):1110–1135, 2010.
13. A. Budhiraja and K. Ross. Convergent numerical scheme for singular stochastic control with state constraints in a portfolio selection problem. *SIAM Journal on Control and Optimization*, 45(6):2169–2206, 2007.
14. R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, first edition, 2004.
15. M.G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
16. M. Dai and F. Yi. Finite-horizon optimal investment with transaction costs: A parabolic double obstacle problem. *Journal of Differential Equations*, 246(4):1445–1469, 2009.

17. M. Dai and Y. Zhong. Penalty methods for continuous-time portfolio selection with proportional transaction costs. *Journal of Computational Finance (to appear)*, 13(3):1–31, 2010.
18. M.H.A. Davis and A.R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713, 1990.
19. P.J. Davis and P. Rabinowitz. *Methods of Numerical Integration.* Dover Publications, second edition, 2007.
20. R.M. Dudley. *Real analysis and probability.* Cambridge University Press, second edition, 2002.
21. E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1(3):281–299, 1995.
22. D.J. Estep. *Practical analysis in one variable.* Springer, 2002.
23. S.N. Ethier and T.G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley, 1986.
24. L.C. Evans. *Partial Differential Equations.* American Mathematical Society, first edition, 1998.
25. W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer-Verlag New York, first edition, 1993.
26. N.C. Framstad, B. Øksendal, and A. Sulem. Optimal consumption and portfolio in a jump diffusion market. In A. Shiryaev and A. Sulem, editors, *Workshop on Mathematical Finance*, INRIA, Paris, 1998.
27. N.C. Framstad, B. Øksendal, and A. Sulem. Optimal consumption and portfolio in a jump diffusion market with proportional transaction costs. *Journal of Mathematical Economics*, 35(2):233–257, 2001.
28. S.A. Gaal. *Point set topology.* Academic Press Inc, 1964.
29. S.R. Ghorpade and B.V. Limaye. *A Course in Multivariable Calculus and Analysis.* Springer, 2010.
30. P.R. Hamos. *Measure Theory.* Springer, 1974.
31. A. Hindy and C.F. Huang. Optimal consumption and portfolio rules with durability and local substitution. *Econometrica*, 61(1):85–121, 1993.
32. Y. Huang and P.A. Forsyth. Analysis of a penalty method for pricing a guaranteed minimum withdrawal beenfit (GMWB). *IMA Journal on Numerical Analysis*, 2009. to appear.
33. Y. Ishikawa. Optimal control problem associated with jump. *Appl. Math. Optim.*, 50(1):21–65, 2004.
34. V.I. Istrăţwscu. *Fixed Point Theory, An Introduction.* Mathematics and its applications. D Reidel Publishing Company, 1981.
35. E.R. Jakobsen and K.H. Karlsen. A "maximum principle for semicontinuous functions" applicable to integro-partial differentiable equations. *NoDEA Nonlinear differential equations and applications*, 13:137–165, 2006.
36. T. Lyche. Lecture notes for inf-mat 4350. Technical report, University of Oslo, 2010.
37. R.C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, 1971.
38. B.F. Nielsen, O. Skavhaug, and A. Tveito. Penalty methods for the numerical solution of American multi-asset option problems. *Journal of Computational and Applied Mathematics*, 222(1):3–16, 2008.
39. T.H. Rydberg. The normal inverse Gaussian Lévy process: Simulation and approximation. *Commun. Statist.-Stochastic Models*, 13(4):887–910, 1997.
40. H.M. Soner. Optimal control with state-space constraint. I. *SIAM J. Control Optim.*, 24(3):552–561, 1986.
41. H.M. Soner. Optimal control with state-space constraint. II. *SIAM J. Control Optim.*, 24(6):1110–1122, 1986.

42. M.I. Taksar and S. Sethi. A note on Merton's 'optimum consumption and portfolio rules in a continuous-time model'. *Journal of Economic Theory*, 46(2):395–401, 1988.
43. J.H. Witte and C. Reisinger. A penalty method for the numerical solution of Hamilton-Jacobi-Bellman (HJB) equations in finance. *Quantitative Finance*, 2010.
44. B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, sixth edition, 2003.

Appendix A
MATLAB code

This appendix contains the most important MATLAB functions used in the simulations in Chapter 15. Each section contains the relevant MATLAB functions for each of three implementations:

(1) implementation of the two-dimensional finite difference scheme described in Chapter 11,
(2) implementation of case A' for the one-dimensional finite difference scheme described in Chapter 13, see Section 15.6, and
(3) implementation of case B' for the one-dimensional finite difference scheme described in Chapter 13, see Section 15.6.

For each of the three implementations, we have the following functions:

(1) the main function $explicit2D/explicit1D\_caseA/explicit1D\_caseB$ that starts the simulation and performs the iteration over the time steps, and
(2) two functions $x0A/x0B/r0$ and $tTA/tTB/tT$ giving, respectively, the boundary condition for $x = 0$ or $r = 0$, and the terminal condition for $t = T$.

Only the one-dimensional programs allow for discontinuous Lévy processes. For the one-dimensional programs, we have some functions associated with the calculation of the integral part of the HJB equation, in addition to the functions described above. For the two-dimensional scheme, we have the utility function U in addition to the functions described above.

The two-dimensional program is relatively generic, as one easily can change the parameter values and the definition of U and W. The one-dimensional scheme for case A' is also relatively generic, as it is easy to change the parameter values, the utility functions and the Lévy measure ν, but, due to the technique of truncating the sum (11.10), the density of ν should be largest close to 0. The program for case B' is specifically made to solve case B', and the only input the user easily can change, are the constants specified in Table 15.6, i.e., not the form of the utility functions or ν.

A.1 Explicit two-dimensional scheme

This section contains code for an explicit version of the two-dimensional scheme described in Section 11, where it is assumed that $\nu \equiv 0$. The input parameters of the program are as follows:

(1) N_x, N_y: the number of grid points in x direction and y direction, respectively,
(2) x_{max}, y_{max}: the maximum value of x and y, respectively,
(3) T: the terminal time,
(4) ϵ: the penalty parameter,
(5) *parameters*: a struct specifying the values of $\hat{\mu}$, σ, $\hat{r}$, δ, γ, β and D.
(6) *BC*: a struct giving the boundary conditions for $y = 0$, $x = x_{max}$ and $y = y_{max}$. In addition to the boundary conditions described in Section 11.3, the Dirichlet boundary condition $V = 0$ is implemented,
(7) *output*: a struct specifying which plots that should be made, and
(8) *refCase*: equal to 1 if W is on the form (15.2) (case A), equal to 2 if W is on the form (7.33) (case B).

The program produces the plots mentioned in the introduction to Chapter 15. If we consider case B ($refCase = 2$), the exact solution formula (7.27) is used to calculate the error of the scheme.

```
% calculates the value function and optimal controls for cases A and B, using
    a two-dimensional scheme
function explicit2D(Nx,Xmax,Ny,Ymax,T,epsilon,parameters,bc,output,refCase)

piFactor = 1; % implies how large fraction of the calculated pi values that
    should be saved. A large piFactor implies that only a small fraction of
    the values are saved
GFactor = 1; % same as piFactor, but for G.

% extracting information from the input parameters
countpi = output.countpi;
countG = output.countG;

y0BC = bc.y0BC;
xinftyBC = bc.xinftyBC;
yinftyBC = bc.yinftyBC;

muhat = parameters.muhat;
sigma = parameters.sigma;
rhat = parameters.rhat;
delta = parameters.delta;
gamma = parameters.gamma;
beta = parameters.beta;
D = parameters.D;
D = D*exp(-delta*T);

% calculating values of rho, pi^*, k, k1, k2, k3 and k4 for the explicit
    solution formula corresponding to case B
rho_a = 2*sigma^2*(rhat+beta);
rho_b = -(2*sigma^2*(rhat+beta)+(muhat-rhat)^2+2*(delta+beta*gamma)*sigma^2);
rho_c = 2*(delta+beta*gamma)*sigma^2;
rho = (-rho_b - sqrt(rho_b^2-4*rho_a*rho_c))/(2*rho_a);
k1 = 1/(gamma*(delta+beta*gamma));
k2 = (1-rho)/((rho-gamma)*(delta+beta*gamma));
```

```
k3 = rho*(1-gamma)/(gamma*(rho-gamma)*(delta+beta*gamma));
k = (1-rho)/(beta*(rho-gamma));
pistar = (muhat-rhat)/(sigma^2*(1-rho));

dx = Xmax/Nx;
dy = Ymax/Ny;

% calculating the smallest value of dt that gives us a stable scheme
dt = 1/(beta*Ymax/dy + muhat*Xmax/dx + (sigma*Xmax)^2/dx^2 + beta/(epsilon*dy
    )+1/(epsilon*dx));
Nt = ceil(T/dt);
dt = T/Nt;

v_old = zeros(Nx+1,Ny+1); % the value of V(t_{m},x,y) for all x, y at a given
     time step
v_new = zeros(Nx+1,Ny+1); % the value of V(t_{m-1},x,y) for all x, y at a
    given time step

% calculating the terminal value of V
if refCase == 1
    for i = 1:(Nx+1)
        for j = 1:(Ny+1)
            v_old(i,j) = tTA((i-1)*dx,(j-1)*dy,gamma,beta,D);
        end
    end
else
    for i = 1:(Nx+1)
        for j = 1:(Ny+1)
            v_old(i,j) = tTB(T,(i-1)*dx,(j-1)*dy,k,k1,k2,k3,delta,gamma,beta,
                rho);
        end
    end
end

% making a vector 'picount' where the calculated values of pi are saved
if countpi == 1
    numberOfPiValues = ceil((Nx-1)*(Ny-1)/piFactor)*Nt;
    picount = zeros(numberOfPiValues,1);
end

% making two matrices x and y containing all relevant x and y values
xvec = linspace(dx,Xmax-dx,Nx-1)';
yvec = linspace(dy,Ymax-dy,Ny-1);
x = zeros(Nx-1,Ny-1);
y = zeros(Nx-1,Ny-1);
for i = 1:(Nx-1)
    for j = 1:(Ny-1)
        x(i,j) = xvec(i);
        y(i,j) = yvec(j);
    end
end

% making a vector 'Gcount' of 0's and 1's, where information about the sign
    of G is saved. The vector 'Rcount' contains the value of x/(x+y)
    corresponding to each element of 'Gcount'
if countG == 1
    numberOfGvalues = ceil((Nx-1)*(Ny-1)/GFactor)*Nt;
    Gcount = zeros(numberOfGvalues,1);
    ratio = x./(x+y);
    Rcount0 = ratio(:);
    Rcount = zeros(ceil((Nx-1)*(Ny-1)/GFactor)*Nt,1);
end

% v_t will contain the simulated value of V(t,xmax/2,ymax/2) for all t
v_t = zeros(1,Nt+1);
```

```
v_t(end) = v_old(ceil(end/2),ceil(end/2));

for m = (Nt+1):-1:2 % iterating over time
    t = (m-1)*dt; % the value of t in the current time step

    % calculating the optimal value of pi in (-infty, infty)
    pi_a = 0.5*sigma^2*x.^2.*dt.*(v_old(3:end,2:(end-1))-2*v_old(2:(end-1)
        ,2:(end-1))+v_old(1:(end-2),2:(end-1)))/dx^2;
    pi_b = (muhat-rhat)*x*dt.*(v_old(3:end,2:(end-1))-v_old(2:(end-1),2:(end
        -1)))/dx;
    pi_c = rhat*x*dt.*(v_old(3:end,2:(end-1))-v_old(2:(end-1),2:(end-1)))/dx;
    pim = -pi_b./(2*pi_a);

    % calculating the optimal value of pi in [0,1]
    % pi_exp = the optimal value of the part of the HJB equation containing
        pi
    pi0_exp = pi_c;
    pi1_exp = pi_a+pi_b+pi_c;
    pim_exp = pi_a.*pim.^2+pi_b.*pim+pi_c;
    pi_exp = zeros(size(pim));
    index = (pim <= 0 | pim >= 1) & pi0_exp >= pi1_exp;
    pi_exp(index) = pi0_exp(index);
    index = (pim <= 0 | pim >= 1) & pi1_exp >= pi0_exp;
    pi_exp(index) = pi1_exp(index);
    index = (pim > 0 & pim < 1 & pim_exp > pi0_exp & pim_exp > pi1_exp);
    pi_exp(index) = pim_exp(index);
    index = (pim > 0 & pim < 1 & pi0_exp > pim_exp & pi0_exp > pi1_exp);
    pi_exp(index) = pi0_exp(index);
    index = (pim > 0 & pim < 1 & pi1_exp > pi0_exp & pi1_exp > pim_exp);
    pi_exp(index) = pi1_exp(index);

    % calculating max{G(D_X V); 0}/epsilon
    Gexp = (dt*beta*(v_old(2:(end-1),3:end) - v_old(2:(end-1),2:(end-1)))/dy
        - dt*(v_old(2:(end-1),2:(end-1))-v_old(1:(end-2),2:(end-1)))/dx)/
        epsilon;
    Gindex = Gexp < 0;
    Gexp(Gindex) = 0;

    % we record the frequency of G>0 for the different values of x/(x+y)
    if countG == 1
        start = ceil((Nx-1)*(Ny-1)/GFactor)*(m-2)+1;
        stop = ceil((Nx-1)*(Ny-1)/GFactor)*(m-1);
        Gindex2 = round(rand(1)*GFactor + linspace(1,length(Gindex(:))-
            GFactor,stop-start+1)); % deciding which grid points we will
            consider
        Gcount(start:stop) = -Gindex(Gindex2)+1; % Gcount(i)=1 corresponds to
             G(D_X V)>0
        Rcount(start:stop) = Rcount0(Gindex2);
    end

    % we record the optimal values of pi
    if countpi == 1
        start = ceil((Nx-1)*(Ny-1)/piFactor)*(m-2)+1;
        stop = ceil((Nx-1)*(Ny-1)/piFactor)*(m-1);
        pi_count0 = pim(Gindex); % we only record values of pi corresponding
            to no consumption
        pi_count0 = pi_count0(round( rand(1)*piFactor +linspace(1,length(
            pi_count0(:))-piFactor,stop-start+1) ));
        picount(start:stop) = pi_count0;
    end

    % calculating the value of V in the interior of the domain
    v_new(2:(end-1),2:(end-1)) = v_old(2:(end-1),2:(end-1)) + dt*exp(-delta*t
        ).*U(y,gamma) - beta*y*dt.*(v_old(2:(end-1),2:(end-1))-v_old(2:(end
        -1),1:(end-2)))/dy + pi_exp + Gexp;
```

```
v_t(m-1) = v_new(ceil(end/2),ceil(end/2));

% boundary condition for x = 0
if refCase == 1
    v_new(1,:) = x0A(t,[0 yvec Ymax],gamma,beta,delta,T,D);
else
    v_new(1,:) = x0B(t,[0 yvec Ymax],k1,gamma,delta);
end

% boundary condition for (x,y)=(Xmax,Ymax)
v_new(end,end) = v_old(end-1,end-1);

% boundary condition for y = 0
% y0BC = 1: assuming G(D_X V) = 0
% y0BC = 2: assuming V_t + F(t,X,D_XV,D_X^2V,I) + max(G(D_X V); 0)/
    epsilon = 0
% y0BC = 3: assuming V_t + F(t,X,D_XV,D_X^2V,I) = 0
if y0BC == 1
    for i = 2:(Nx+1)
        v_new(i,1) = (beta*dx*v_new(i,2) + dy*v_new(i-1,1))/(dy+dx*beta);
    end
else
    % we calculate the terms of the HJB equation as for interior points
    yval = 0;
    pi_a = 0.5*sigma^2*xvec.^2.*dt.*(v_old(3:end,1)-2*v_old(2:(end-1),1)+
        v_old(1:(end-2),1))/dx^2;
    pi_b = (muhat-rhat).*xvec.*dt.*(v_old(3:end,1)-v_old(2:(end-1),1))/dx
        ;
    pi_c = rhat*xvec*dt.*(v_old(3:end,1)-v_old(2:(end-1),1))/dx;
    pim = -pi_b./(2*pi_a);
    pi0_exp = pi_c;
    pi1_exp = pi_a+pi_b+pi_c;
    pim_exp = pi_a.*pim.^2+pi_b.*pim+pi_c;
    pi_exp = zeros(size(pim));
    index = (pim <= 0 | pim >= 1) & pi0_exp >= pi1_exp;
    pi_exp(index) = pi0_exp(index);
    index = (pim <= 0 | pim >= 1) & pi1_exp >= pi0_exp;
    pi_exp(index) = pi1_exp(index);
    index = (pim > 0 & pim < 1 & pim_exp > pi0_exp & pim_exp > pi1_exp);
    pi_exp(index) = pim_exp(index);
    index = (pim > 0 & pim < 1 & pi0_exp > pim_exp & pi0_exp > pi1_exp);
    pi_exp(index) = pi0_exp(index);
    index = (pim > 0 & pim < 1 & pi1_exp > pi0_exp & pi1_exp > pim_exp);
    pi_exp(index) = pi1_exp(index);
    Gexp = (dt*beta*(v_old(2:(end-1),2)-v_old(2:(end-1),1))/dy - dt*(
        v_old(2:(end-1),1)-v_old(1:(end-2),1))/dx)/epsilon;
    if y0BC == 2
        index = Gexp < 0;
        Gexp(index) = 0;
    else
        Gexp = Gexp*0;
    end
    v_new(2:(end-1),1) = v_old(2:(end-1),1) + dt*exp(-delta*t).*U(yval,
        gamma) + pi_exp + Gexp;
end

% boudary condition for x = Xmax
if xinftyBC == 1 % assuming G(D_X V) = 0
    for j=Ny:-1:1
        v_new(Nx+1,j) = (beta*dx*v_new(Nx+1,j+1)+dy*v_new(Nx,j))/(dy+dx*
            beta);
    end
elseif xinftyBC == 2 % Neuman boundary condition
    v_new(Nx+1,1:(end-1)) = v_new(Nx,1:(end-1));
```

```
    elseif xinftyBC == 3 % Dirichlet boundary condition
        v_new(Nx+1,1:(end-1)) = 0;
    end

    % boundary condition for y = Ymax
    if yinftyBC == 1 % assuming V_t + F(t,X,D_XV,D_X^2V,I) = 0
        yval = Ymax;
        pi_a = 0.5*sigma^2.*xvec.^2*dt.*(v_old(3:end,end) - 2*v_old(2:(end-1)
            ,end)+v_old(1:(end-2),end))/dx^2;
        pi_b = (muhat-rhat)*xvec*dt.*(v_old(3:end,end)-v_old(2:(end-1),end))/
            dx;
        pi_c = rhat*xvec*dt.*(v_old(3:end,end)-v_old(2:(end-1),end))/dx;
        pim = -pi_b./(2*pi_a);
        pi0_exp = pi_c;
        pi1_exp = pi_a+pi_b+pi_c;
        pim_exp = pi_a.*pim.^2+pi_b.*pim+pi_c;
        pi_exp = 0*pim;
        index = (pim <= 0 | pim >= 1) & pi0_exp >= pi1_exp;
        pi_exp(index) = pi0_exp(index);
        index = (pim <= 0 | pim >= 1) & pi1_exp >= pi0_exp;
        pi_exp(index) = pi1_exp(index);
        index = (pim > 0 & pim < 1 & pim_exp > pi0_exp & pim_exp > pi1_exp);
        pi_exp(index) = pim_exp(index);
        index = (pim > 0 & pim < 1 & pi0_exp > pim_exp & pi0_exp > pi1_exp);
        pi_exp(index) = pi0_exp(index);
        index = (pim > 0 & pim < 1 & pi1_exp > pi0_exp & pi1_exp > pim_exp);
        pi_exp(index) = pi1_exp(index);
        v_new(2:(end-1),end) = v_old(2:(end-1),end) + dt*exp(-delta*t).*U(
            yval,gamma) - beta*yval*dt.*(v_old(2:(end-1),end)-v_old(2:(end-1)
            ,end-1))/dy + pi_exp;
    elseif yinftyBC == 2 % Neuman boundary condition
        v_new(2:(end-1),end) = v_new(2:(end-1),end-1);
    elseif yinftyBC == 3 % Dirichlet boundary condition
        v_new(2:(end-1),end) = 0*v_new(2:(end-1),end);
    else % ignoring the increase in Y due to consumption
        a1 = Ymax.^gamma.*exp(beta*gamma*t)/gamma;
        a2 = delta+beta*gamma;
        v_new(2:(end-1),end) = a1.*exp(-a2*t)/a2 - a1.*exp(-a2*T)/a2 + D*exp(
            beta*gamma*(t-T))*(Ymax+beta*xvec*exp(rhat*(T-t))).^gamma/gamma;
    end

    % the solution found in this iteration, is input in the next iteration
    v_old = v_new;
end

% we calculate the exact solution and the error if we consider reference case
    B
if refCase == 2
    Vexact = zeros(Nx+1,Ny+1); % the exact solution
    for i=1:(Nx+1)
        for j = 1:(Ny+1)
            x = (i-1)*dx;
            y = (j-1)*dy;
            if x<k*y
                Vexact(i,j) = k1*y^gamma + k2*y^gamma*(x/(k*y))^rho;
            else
                Vexact(i,j) = k3*(y+beta*x)^gamma/(1+beta*k)^gamma;
            end
        end
    end
    error = max(max(abs(Vexact(1:round(0.8*Nx),1:round(0.8*Ny))-v_new(1:round
        (0.8*Nx),1:round(0.8*Ny)))));
    display(['error ' num2str(error)]);
end
```

```
% plotting the value function for t=0
figure;
[xMesh yMesh] = meshgrid(linspace(0,Xmax,Nx+1),linspace(0,Ymax,Ny+1));
surf(xMesh,yMesh,v_new');
xlabel('x','FontSize',12);
ylabel('y','Fontsize',12);
zlabel('V','FontSize',12);
title('Value function for t=0','FontSize',12);
colormap('gray');

% plotting the value function  for t=0, y=Ymax/2
figure;
plot(linspace(0,Xmax,Nx+1),v_new(:,ceil(end/2)),'k--');
xlabel('x','FontSize',12);
ylabel('V','Fontsize',12);
title(['Value function for y=' num2str(Ymax/2)],'FontSize',12);
hold on;
if refCase == 2
    plot(linspace(0,Xmax,length(Vexact(:,ceil(end/2)))),Vexact(:,ceil(end/2))
        ,'k--');
end

% plotting the distribution of the calculated values of pi
if countpi == 1
    index = picount>2;
    picount(index)=2; % 'pi'-values above 2 are registered as 2
    index = picount<0;
    picount(index) = 0; % 'pi'-values below 0 are registered as 0
    figure;
    hist(picount,500);
    xlabel('\pi','FontSize',12)
    ylabel('frequency','FontSize',12)
    title('Distribution of \pi for t\in [0,T]')
end

% plotting the relative frequency of G>0 as a function of x/(x+y)
if countG == 1
    figure;
    antBins = 200; % number of intervals
    bins0 = linspace(0,1,antBins+1);
    bins = bins0(1:antBins);
    maxFreq = hist(Rcount,bins);
    Rcount_all = Rcount(Gcount == 1);
    freq = hist(Rcount_all,bins); % frequency of G>0
    frac = freq./maxFreq; % relative frequency of G>0
    binsPlot = (bins0(1:antBins)+bins0(2:(antBins+1)))/2;
    hold on;
    plot(binsPlot,frac,'k--');
    xlabel('x/(x+y)','FontSize',12);
    ylabel('relative frequency of G>0','FontSize',12)
    title('Relative frequency of G>0 for t\in[0,T]');
end
end
```

```
% the utility function
function u = U(y,gamma)
u = y.^gamma./gamma;
end
```

```
% boundary condition for x=0, case A
function answer = x0A(t,y,gamma,beta,delta,T,D)
a1 = y.^gamma.*exp(beta*gamma*t)/gamma;
a2 = delta+beta*gamma;
a3 = exp(-a2*t)/a2 - exp(-a2*T)/a2 + D*exp(-beta*gamma*T);
answer = a1.*a3;
end

% boundary condition for x=0, case B
function answer = x0B(t,y,k1,gamma,delta)
answer = exp(-delta*t).*k1.*y.^gamma;
end
```

```
% terminal condition, case A
function answer = tTA(x,y,gamma,beta,D)
answer = D*(y+beta*x).^gamma/gamma;
end

% terminal condition, case B
function answer = tTB(T,x,y,k,k1,k2,k3,delta,gamma,beta,rho)
if x < k*y
    answer = k1*y^gamma + k2*y^gamma*(x/(k*y))^rho;
else
    answer = k3*(y+beta*x)^gamma/(1+beta*k)^gamma;
end
answer = exp(-delta*T)*answer;
end
```

A.2 Explicit one-dimensional scheme, case A'

This section contains the MATLAB code for case A'. The part of the code where graphs are made, is very similar to the corresponding code in Appendix A.1 above, and is therefore omitted. We have also omitted the part of the code that record the value of π and the sign of $G(D_X V)$.

As mentioned above, this implementation allows for discontinuous Lévy processes. This implies that an integral must be calculated numerically at each time step, and these functions are added to the program due to the integral calculations:

(1) *calcJump_caseA*: calculates the integral part of the HJB equation at each time step,
(2) *kintint*: calculates $\tilde{k}$, defined by (11.9),
(3) *bTildeIntegrand*: calculates the integrand of the integral defining $\tilde{b}$, divided by πx, see Section 11.2.1, and
(4) *eta*: the functions η given by (11.2).

The functions *kintint* and *bTildeIntegrand* are used in time consuming integral calculations, and they are *not* called at each time step, only in the beginning of *explicit1D_caseA*.

The input parameters of the main function *explicit1D_caseA* are as follows:

(1) N_r: the number of points in the r direction,
(2) r_{max}: the maximum value of the ratio $r = x/y$,
(3) T: the terminal time,
(4) ϵ: the penalty parameter,
(5) *parameters*: a struct specifying the value of $\hat{\mu}$, σ, $\hat{r}$, δ, γ, β, D and b, and
(6) *deltaPi* : the maximum difference between π for t_{m-1} and t_m, $m \in \{1, \ldots, N_t\}$, see Section 14.

The input parameters of the function *calcJump_caseA*, calculating the integral part of the HJB equation, are as follows

(1) π: a vector of π-values for interior grid points (length: $N_r - 1$),
(2) r_{max}: the maximum value of the ratio $r = x/y$,
(3) V: a vector giving the value function at all grid points for the current time step (length: $N_r + 1$),
(4) $N_{max,M}$, $N_{max,P}$: the number of terms in the sum (11.10) for negative and positive z, respectively
(5) *kTildeM*, *kTildeP*: vectors containing, respectively, the value of $\tilde{k}^-_{\mathcal{H},n}$ for $n = 0, \ldots N_{max,M}$, and the value of $\tilde{k}^+_{\mathcal{H},n}$ for $n = 0, \ldots, N_{max,P}$.

```
% calculates the value function for case A', using a one-dimensional scheme
function explicit1D_caseA(Nr,Rmax,Tmax,epsilon,parameters,piInitial,deltaPi)

% extracing information from the input parameters
muhat=parameters.muhat;
sigma=parameters.sigma;
rhat=parameters.rhat;
gamma=parameters.gamma;
beta=parameters.beta;
lambda=parameters.lambda;
D=parameters.D;
delta=parameters.delta;

D=D*exp(-delta*Tmax);

% calculating dt
dr=Rmax/Nr;
dtlimit = 1/(0.5*sigma*Rmax^2*2/dr^2 + (muhat-rhat)*Rmax/dr + rhat*Rmax/dr +
    (beta*Rmax+1)/(dr*epsilon) - beta*gamma/epsilon + beta*gamma + beta*Rmax/
    dr);
Nt=ceil(Tmax/dtlimit);
dt=Tmax/Nt;

% V_new contains the calculated value function at each time step
V_new=zeros(1,Nr+1);

% the values r in our grid
rVect=linspace(dr,Rmax-dr,Nr-1);

% calculating the terminal value of V
V_old=tT([0 rVect Rmax],gamma,beta,D);

% At each time step, pi_old is the values of pi used in the previous step of
    the interation
pi_old= piInitial*ones(size(V_new(2:(end-1))));
piTol=deltaPi;
```

```
% We calculate some integrals needed in the calculation of the integral part
    of the HJB equation
NmaxM=100; % the maximum number of terms in the sum approximating the
    integral for negative z
NmaxP=100; % same as NmaxM, but for positive z
kTildeM = zeros(NmaxM+1,1); % integrals for negative z
kTildeP = zeros(NmaxP+1,1); % integrals for positive z
tol=0.0001; % terms that are less than 'tol' of the largest term, are omitted
for n=1:(NmaxM+1) % calculating integrals for negative z
    kTildeM(n) = quadl(@kintint,-(n)*dx,-(n-1)*dx);
    if abs(kTildeM(n))<tol*max(abs(kTildeM))
        NmaxM = n-2;
        break;
    end
end
for n=1:(NmaxP+1) % calculating integrals for positive z
    kTildeP(n) = quadl(@kintint,(n-1)*dx,(n)*dx);
    if abs(kTildeP(n))<tol*max(abs(kTildeP))
        NmaxP = n-2;
        break;
    end
end
kTildeM=kTildeM(1:(NmaxM+1));
kTildeP=kTildeP(1:(NmaxP+1));
bTilde = (quadl(@bTildeIntegrand,-500,-100) + quadl(@bTildeIntegrand,-100,-1)
     + quadl(@bTildeIntegrand,-1,-0.1) + quadl(@bTildeIntegrand,-0.1,0) +
    quadl(@bTildeIntegrand,0,0.1) + quadl(@bTildeIntegrand,0.1,1) + quadl(
    @bTildeIntegrand,1,10) + quadl(@bTildeIntegrand,10,100) + quadl(
    @bTildeIntegrand,100,500));

for m=Nt:-1:1 % we interate over the time interval [0, T]
    t=m*dt; % the value of t at the current time step

    % calculating the optimal value of pi, assuming the optimal value of pi
        is either pi_old, pi_old-piTol eller pi_old+piTol, where pi_old is
        the value of pi used in the previous time step
    acoeff = 0.5*sigma^2*rVect.^2.*(V_old(3:end)-2*V_old(2:(end-1))+V_old(1:(
        end-2)))/dr^2;
    bcoeff = (muhat-rhat)*rVect.*(V_old(3:end)-V_old(2:(end-1)))/dr;
    ccoeff = rhat*rVect.*(V_old(3:end)-V_old(2:(end-1)))/dr;

    pip=(pi_old+piTol); % increase in pi compared to the previous time step
    pi0=pi_old; % the value of pi used in the previous time step
    pim=(pi_old-piTol); % decrease in pi compared to the previous time step
    index = pip>1; pip(index)=1; % the used value of pi must be <= 1
    index = pim<0; pim(index)=0; % the used value of pi must be >= 0

    % pi_exp = the value of the part of the HJB equation containing pi,
        except for the integral term
    pip_exp = acoeff.*pip.^2 + bcoeff.*pip + ccoeff;
    pi0_exp = acoeff.*pi0.^2 + bcoeff.*pi0 + ccoeff;
    pim_exp = acoeff.*pim.^2 + bcoeff.*pim + ccoeff;

    % calculating max{G(D_X V); 0}/epsilon
    Gexp = -(beta*rVect+1).*(V_old(2:(end-1))-V_old(1:(end-2)))/dr + beta*
        gamma*V_old(2:(end-1));
    Gindex=Gexp<0;
    Gexp(Gindex)=0;

    % calculating the integral part of the HJB equation for each of the three
         pi values we consider
    jumpp = lambda*calcJump_caseA(pip',Rmax,V_old',NmaxM,NmaxP,kTildeM,
        kTildeP,bTilde)';
    jump0 = lambda*calcJump_caseA(pi0',Rmax,V_old',NmaxM,NmaxP,kTildeM,
        kTildeP,bTilde)';
```

```
    jumpm = lambda*calcM10(pim',Rmax,V_old',NmaxM,NmaxP,kTildeM,kTildeP,
        bTilde)';

    % we choose the maximizing value of pi
    maxTerm=max([(jumpp+pip_exp); (jump0+pi0_exp); (jumpm+pim_exp)])';

    % we record the used values of pi, which will be used to determine pi in
        the next iteration
    indexp = (maxTerm==(jumpp+pip_exp)');
    pi_old(indexp')=pip(indexp');
    index0 = (maxTerm==(jump0+pi0_exp)');
    pi_old(index0')=pi0(index0');
    indexm = (maxTerm==(jumpm+pim_exp)');
    pi_old(indexm')=pim(indexm');

    % calculating the value function for interior points of our domain
    V_new(2:(end-1)) = V_old(2:(end-1)) + dt*exp(-delta*t)/gamma + dt*beta.*
        rVect.*(V_old(3:end)-V_old(2:(end-1)))/dr - dt*beta*gamma*V_old(2:(
        end-1)) + dt*Gexp/epsilon + dt*maxTerm';

    % Dirichlet boundary condition at r=0
    V_new(1)=r0(t,gamma,beta,delta,Tmax,D);

    % Assuming G(V,V_r)=0 for r=r_max
    V_new(end)=V_new(end-1)*(beta*Rmax+1)/(beta*Rmax+1-beta*gamma*dr);

    % the solution found in this iteration, is input in the next iteration
    V_old=V_new;
end
end
```

```
% boundary condition for r=0
function answer = r0(t,gamma,beta,delta,T,D)
a1 = exp(beta*gamma*t)/gamma;
a2 = delta+beta*gamma;
a3 = exp(-a2*t)/a2 - exp(-a2*T)/a2 + D*exp(-beta*gamma*T);
answer = a1.*a3;
end
```

```
% terminal condition
function v=tT(r,gamma,beta,D)
v=D*(1+beta*r).^gamma/gamma;
end
```

```
% double integral of the density function k of the Levy measure
function answer=kintint(z)
% k(z) = exp(-0.1|z|)/|z|^2
b=0.1;
length0 = length(z);
answer = zeros(1,length0);
for i=1:length0
    if z(i) > 0
        answer(i) = mfun('Ei',1,b*z(i)) - exp(-b*z(i)) + b*z(i)*mfun('Ei',1,b
            *z(i));
    else
```

```
        answer(i) = mfun('Ei',1,-b*z(i)) - exp(2*b(i)) - 2*b(i)*mfun('Ei',1,-
            b*z(i));
    end
end
end
```

```
% integrand of the integral defining bTilde, divided by pi*x
function answer=bTildeIntegrand(z)
answer=exp(z).*kintint(z);
end
```

```
% calculates the integral part of the HJB equation for case A'
function jump = calcJump_caseA(pi,Xmax,phi_val,NmaxM,NmaxP,kTildeM,kTildeP,
    bTildeIntegral)
Nr=length(pi)+1; % number of grid points minus 1
dr=Xmax/Nr; % delta r
r=linspace(dr,Xmax-dr,Nr-1)'; % vector containing interior values of r

A = zeros(length(r),Nr+1);  % coefficients of each v_i in the calculation of
    the jump
dz = sqrt(dr);
rindd0=round(r/dr);

% For each term of the two sums approximating the integral, we calculate D_X
    ^2v numerically by a three point method. We call the three points 'r_m',
    'r_z' and 'r_p'. Each of the three points 'r_x' (x=m,z,p) lie between two
    points r_i and r_{i+1} of our grid. The vectors 'wpp1', 'wpm1' and 'wpz1
    ' give us the weight of v(r_i) in the linear interpolation where we
    approximate v(r) by v(r_i) and v(r_{i+1}) for the sum with z>0. The
    vectors 'lpm', 'lpz' and 'lpp' denote the index i for r_m, r_z and r_p,
    respectively, for the sum with z>0. The vectors 'wmp1', 'wmm1', 'wmz1', '
    lmm', 'lmz' and 'lmp' give us the same information for the case z<0.
wpp1 = zeros(length(pi),NmaxP+1);
wpm1 = zeros(length(pi),NmaxP+1);
wpz1 = zeros(length(pi),NmaxP+1);
wmp1 = zeros(length(pi),NmaxM+1);
wmm1 = zeros(length(pi),NmaxM+1);
wmz1 = zeros(length(pi),NmaxM+1);

n=1:(NmaxP+1);
[zMesh,sMesh]=meshgrid(n,r);

% calculating 'wpm' and 'lpm'
[zMesh,piMesh]=meshgrid(dr*(n-0.5)-dz,pi);
wpm = eta(sMesh,zMesh,piMesh)/dr;
lpm = floor(wpm);
wpm = wpm - lpm;

% calculating 'wpp' and 'lpp'
[zMesh,piMesh]=meshgrid(dr*(n-0.5)+dz,pi);
wpp=eta(sMesh,zMesh,piMesh)/dr;
lpp = floor(wpp);
wpp = wpp - lpp;

% calculating 'wpz' and 'lpz'
[zMesh,piMesh]=meshgrid(dr*(n-0.5),pi);
wpz = eta(sMesh,zMesh,piMesh)/dr;
lpz = floor(wpz);
```

```
wpz = wpz - lpz;

% removing contributions corresponding to points outside our grid
rindd0Mesh = meshgrid(rindd0,ones(NmaxP+1,1))';
index = (rindd0Mesh+lpp<Nr);
lpm(~index)=0;
wpm(~index)=0;
wpm1(index)=1-wpm(index);
lpp(~index)=0;
wpp(~index)=0;
wpp1(index)=1-wpp(index);
lpz(~index)=0;
wpz(~index)=0;
wpz1(index)=1-wpz(index);

n=1:(NmaxM+1);
[zMesh,sMesh]=meshgrid(n,r);

% calculating 'wmm' and 'lmm'
[zMesh,piMesh]=meshgrid(-dr*(n-0.5)-dz,pi);
wmm = eta(sMesh,zMesh,piMesh)/dr;
lmm = floor(wmm);
wmm = wmm - lmm;

% calculating 'wmp' and 'lmp'
[zMesh,piMesh]=meshgrid(-dr*(n-0.5)+dz,pi);
wmp = eta(sMesh,zMesh,piMesh)/dr;
lmp = floor(wmp);
wmp = wmp - lmp;

% calculating 'wmz' and 'lmz'
[zMesh,piMesh]=meshgrid(-dr*(n-0.5),pi);
wmz = eta(sMesh,zMesh,piMesh)/dr;
lmz = floor(wmz);
wmz = wmz - lmz;

% removing contributions corresponding to points outside our grid
rindd0Mesh = meshgrid(rindd0,ones(NmaxM+1,1))';
index = (rindd0Mesh+lmp+1<Nr & rindd0Mesh+lmm>1);
lmm(~index)=0;
wmm(~index)=0;
wmm1(index)=1-wmm(index);
lmp(~index)=0;
wmp(~index)=0;
wmp1(index)=1-wmp(index);
lmz(~index)=0;
wmz(~index)=0;
wmz1(index)=1-wmz(index);

% adding contributions to the A-matrix from the sum for positive z
kTildePMesh=meshgrid(kTildeP,ones(size(r)));
rindd0Mesh=meshgrid(rindd0,lpm(1,:))';

% adding contributions to the A-matrix from the points 'r_m' in the sum for
    positive z
rindd1 = rindd0Mesh + (lpm+1);
rindd2 = rindd0Mesh + (lpm);
rinddNew1=0*lpm;
rinddNew2=0*lpm;
length_r=length(r);
for i=1:size(lpm,1)
    rinddNew1(i,:) = rindd1(i,:)*length_r+i-1;
    rinddNew2(i,:) = rindd2(i,:)*length_r+i-1;
end
for n=1:(NmaxP+1)
```

```
    A(rinddNew1(:,n)+1) = A(rinddNew1(:,n)+1)+wpm(:,n).*kTildePMesh(:,n)/dz
        ^2;
    A(rinddNew2(:,n)+1) = A(rinddNew2(:,n)+1)+wpm1(:,n).*kTildePMesh(:,n)/dz
        ^2;
end

% adding contributions to the A-matrix from the points 'r_p' in the sum for
    positive z
rindd1 = rindd0Mesh + (lpp+1);
rindd2 = rindd0Mesh + (lpp);
for i=1:size(lpm,1)
    rinddNew1(i,:) = rindd1(i,:)*length_r+i-1;
    rinddNew2(i,:) = rindd2(i,:)*length_r+i-1;
end
for n=1:(NmaxP+1)
    A(rinddNew1(:,n)+1) = A(rinddNew1(:,n)+1)+wpp(:,n).*kTildePMesh(:,n)/dz
        ^2;
    A(rinddNew2(:,n)+1) = A(rinddNew2(:,n)+1)+wpp1(:,n).*kTildePMesh(:,n)/dz
        ^2;
end

% adding contributions to the A-matrix from the points 'r_z' in the sum for
    positive z
rindd1 = rindd0Mesh + (lpz+1);
rindd2 = rindd0Mesh + (lpz);
for i=1:size(lpm,1)
    rinddNew1(i,:) = rindd1(i,:)*length_r+i-1;
    rinddNew2(i,:) = rindd2(i,:)*length_r+i-1;
end
for n=1:(NmaxP+1)
    A(rinddNew1(:,n)+1) = A(rinddNew1(:,n)+1)-2*wpz(:,n).*kTildePMesh(:,n)/dz
        ^2;
    A(rinddNew2(:,n)+1) = A(rinddNew2(:,n)+1)-2*wpz1(:,n).*kTildePMesh(:,n)/
        dz^2;
end

% adding contributions to the A-matrix from the sum for positive z
kTildeMMesh=meshgrid(kTildeM,ones(size(r)));
rindd0Mesh=meshgrid(rindd0,lmm(1,:))';
rinddNew1=zeros(size(lmm));
rinddNew2=zeros(size(lmm));

% adding contributions to the A-matrix from the points 'r_m' in the sum for
    negative z
rindd1 = rindd0Mesh + (lmm+1);
rindd2 = rindd0Mesh + (lmm);
for i=1:size(lmm,1)
    rinddNew1(i,:) = rindd1(i,:)*length_r+i-1;
    rinddNew2(i,:) = rindd2(i,:)*length_r+i-1;
end
for n=1:(NmaxM+1)
    A(rinddNew1(:,n)+1) = A(rinddNew1(:,n)+1)+wmm(:,n).*kTildeMMesh(:,n)/dz
        ^2;
    A(rinddNew2(:,n)+1) = A(rinddNew2(:,n)+1)+wmm1(:,n).*kTildeMMesh(:,n)/dz
        ^2;
end

% adding contributions to the A-matrix from the points 'r_m' in the sum for
    negative z
rindd1 = rindd0Mesh + (lmp+1);
rindd2 = rindd0Mesh + (lmp);
for i=1:size(lmm,1)
    rinddNew1(i,:) = rindd1(i,:)*length_r+i-1;
    rinddNew2(i,:) = rindd2(i,:)*length_r+i-1;
end
```

```
for n=1:(NmaxM+1)
    A(rinddNew1(:,n)+1) = A(rinddNew1(:,n)+1)+wmp(:,n).*kTildeMMesh(:,n)/dz
        ^2;
    A(rinddNew2(:,n)+1) = A(rinddNew2(:,n)+1)+wmp1(:,n).*kTildeMMesh(:,n)/dz
        ^2;
end

% adding contributions to the A-matrix from the points 'r_m' in the sum for
    negative z
rindd1 = rindd0Mesh + (lmz+1);
rindd2 = rindd0Mesh + (lmz);
for i=1:size(lmm,1)
    rinddNew1(i,:) = rindd1(i,:)*length_r+i-1;
    rinddNew2(i,:) = rindd2(i,:)*length_r+i-1;
end
for n=1:(NmaxM+1)
    A(rinddNew1(:,n)+1) = A(rinddNew1(:,n)+1)-2*wmz(:,n).*kTildeMMesh(:,n)/dz
        ^2;
    A(rinddNew2(:,n)+1) = A(rinddNew2(:,n)+1)-2*wmz1(:,n).*kTildeMMesh(:,n)/
        dz^2;
end

% jump = the value of the integral for each grid interior point r_i
jump=A*phi_val;
bTilde = r.*pi.*bTildeIntegral;
jump = jump - bTilde.*(phi_val(2:(end-1),:)-phi_val(1:(end-2),:))/dr;
end
```

```
% the change in wealth due to a jump of size z
function e=eta(r,z,pi)
e = pi.*r.*(exp(z)-1);
end
```

A.3 Explicit one-dimensional scheme, case B'

This section contains the MATLAB code for case B'. As for case A', we have omitted the part of the code where the graphs are made, and the part of the code where the π values and the sign of $G(D_X V)$ are saved. The form of the Lévy measure ν is different for case B' compared to case A', and therefore the part of the program associated with the integral calculation is somewhat different:

(1) The calculation of the integrals $\tilde{k}^{\pm}_{\mathcal{H},n}$ and $\tilde{b}^a$ can be done analytically for case B'. The integral calculations are performed in the function *kintint*, incorporated in the rest of the code, instead of being calculated in the beginning of the program.
(2) Since $\nu\big((0,\infty)\big) = 0$ for case B', we do not have to calculate $\mathcal{I}^{a,+}_{\mathcal{H}}$, only $\mathcal{I}^{a,-}_{\mathcal{H}}$.
(3) The sum (11.10) consists of a finite number of terms for case B', and therefore we do not need to truncate the sum.

In addition to the functions *explicit1D_caseB*, *r0B* and *tTB*, the program consists of the following functions:

(1) *calcJump_caseB*: calculates the integral part of the HJB equation at each time step, and
(2) *calcDeltaAndRho*: solves the system of equations (7.6) (called (E) in the program) with δ and ρ as unknowns.

The input parameters to *explicit1D_caseB* are the same as for case A', except for the input variable *parameters*. The optimal value of π (π^*) and ξ are input parameters to the program instead of δ and b. As explained in Chapter 14, the reason for having π^* as an input parameter instead of δ, is that the system of equations (7.6) is difficult to solve with ρ and π^* as unknowns, while it is easy to solve with ρ and δ as unknowns.

The input parameters to *calcJump_caseB* are the same as for *calcJump_caseA*, except that $N_{max,M}$, $N_{max,P}$, *kTildeM* and *kTildeP* are not given as input, but calculated inside the function.

```
% calculates the value function for case B', using a one-dimensional scheme
function explicit1D_caseB(Nr,Rmax,Tmax,epsilon,parameters,deltaPi)

% extracting information from the input parameters
muhat=parameters.muhat;
sigma=parameters.sigma;
rhat=parameters.rhat;
piInitial=parameters.pi0;
gamma=parameters.gamma;
beta=parameters.beta;
xsi=parameters.xsi;
lambda=parameters.lambda;

% solving the system of equations (E)
[delta rho]=calcDeltaAndRho(parameters);

% calculating the constants of the exact solution
k1 = 1/(gamma*(delta+beta*gamma));
k2 = (1-rho)/((rho-gamma)*(delta+beta*gamma));
k3 = rho*(1-gamma)/(gamma*(rho-gamma)*(delta+beta*gamma));
k = (1-rho)/(beta*(rho-gamma));

dr=Rmax/Nr;
dtlimit = 1/(0.5*sigma*Rmax^2*2/dr^2 + (muhat-rhat)*Rmax/dr + rhat*Rmax/dr +
    (beta*Rmax+1)/(dr*epsilon) - beta*gamma/epsilon + beta*gamma + beta*Rmax/
    dr);
Nt=ceil(Tmax/dtlimit);
dt=Tmax/Nt;

V_new=zeros(1,Nr+1); % the value function
V_old=tT(Tmax,linspace(0,Rmax,Nr+1),k,k1,k2,k3,delta,gamma,beta,rho); %
    terminal condition
rVect=linspace(dr,Rmax-dr,Nr-1);

% calculate pi_old
pi_old= piInitial*ones(size(V_new(2:(end-1))));
piTol=deltaPi;

for m=Nt:-1:1  % iterating over time
    t=m*dt; % the value of pi at the current time step

    % calculating the optimal value of pi, assuming the optimal value of pi
        is either pi_old, pi_old-piTol eller pi_old+piTol, where pi_old is
        the value of pi used in the previous time step
    acoeff = 0.5*sigma^2*rVect.^2.*(V_old(3:end)-2*V_old(2:(end-1))+V_old(1:(
        end-2)))/dr^2;
```

```
    bcoeff = (muhat-rhat)*rVect.*(V_old(3:end)-V_old(2:(end-1)))/dr;
    ccoeff = rhat*rVect.*(V_old(3:end)-V_old(2:(end-1)))/dr;

    pip=pi_old+piTol; % increase in pi compared to the previous time step
    pi0=pi_old; % the value of pi used in the previous time step
    pim=pi_old-piTol; % decrease in pi compared to the previous time step
    index = pip>1; pip(index)=1; % the used value of pi must be <= 1
    index = pim<0; pim(index)=0; % the used value of pi must be >= 0

    % pi_exp = the value of the part of the HJB equation containing pi,
        except for the integral term
    pip_exp = acoeff.*pip.^2 + bcoeff.*pip + ccoeff;
    pi0_exp = acoeff.*pi0.^2 + bcoeff.*pi0 + ccoeff;
    pim_exp = acoeff.*pim.^2 + bcoeff.*pim + ccoeff;

    % calculating max{G(D_X V); 0}/epsilon
    Gexp = -(beta*rVect+1).*(V_old(2:(end-1))-V_old(1:(end-2)))/dr + beta*
        gamma*V_old(2:(end-1));
    Gindex=Gexp<0;
    Gexp(Gindex)=0;

    % calculating the integral part of the HJB equation for each of the three
        pi values we consider
    jumpp = lambda*calcJump_caseB(pip',Rmax,V_old',xsi)';
    jump0 = lambda*calcJump_caseB(pi0',Rmax,V_old',xsi)';
    jumpm = lambda*calcJump_caseB(pim',Rmax,V_old',xsi)';

    % we choose the maximizing value of pi
    maxTerm=max([(jumpp+pip_exp); (jump0+pi0_exp); (jumpm+pim_exp)])';

    % we record the used values of pi, which will be used to determine pi in
        the next iteration
    indexp = (maxTerm==(jumpp+pip_exp)');
    pi_old(indexp')=pip(indexp');
    index0 = (maxTerm==(jump0+pi0_exp)');
    pi_old(index0')=pi0(index0');
    indexm = (maxTerm==(jumpm+pim_exp)');
    pi_old(indexm')=pim(indexm');

    % calculating the value function for interior points of our domain
    V_new(2:(end-1)) = V_old(2:(end-1)) + dt*exp(-delta*t)/gamma + dt*beta.*
        rVect.*(V_old(3:end)-V_old(2:(end-1)))/dr - dt*beta*gamma*V_old(2:(
        end-1)) + dt*Gexp/epsilon + dt*maxTerm';

    % Dirichlet boundary condition at r=0
    V_new(1)=r0(t,k1,delta);

    % Assuming G(V,V_r)=0 for r=r_max
    V_new(end)=V_new(end-1)*(beta*Rmax+1)/(beta*Rmax+1-beta*gamma*dr);

    % the solution found in this iteration, is input in the next iteration
    V_old=V_new;
end
end
```

```
% solving the system of equations (E) with delta and rho as unknowns
function [delta rho]=calcDeltaAndRho(params)

% extracting the input parameters
pi0=params.pi0;
muhat=params.muhat;
rhat=params.rhat;
```

```
sigma=params.sigma;
gamma=params.gamma;
beta=params.beta;
xsi = params.xsi;
lambda=params.lambda;

% finding the root of the first equation of (E)
eta=1-exp(-xsi);
exp1 = 1-pi0*eta;
rhos=linspace(gamma,1,10000);
exp2=(-sigma^2*(1-rhos)*pi0+(muhat-rhat))/(lambda*eta)+1;
y_term = log(exp2)/log(exp1)+1-rhos;
for i=1:(length(y_term)-1)
    if y_term(i)*y_term(i+1) <= 0;
        break;
    end
end

% if the first equation of (E) do not have any roots, the parameters are not
    valid
if i==length(y_term)-1
    display('not valid choice of paramers')
    delta=-100; rho=-100;
    return;
end

% Our value of rho is a linear combination of rhos(i) and rhos(i+1). If we
    let LHS denote the left-hand side of the first equation of (E), rho=rhos(
    i) implies LHS <= RHS, while rho=rhos(i+1) implies LHS >= RHS.
rho= (rhos(i)*abs(y_term(i+1)) + rhos(i+1)*abs(y_term(i)))/(abs(y_term(i))+
    abs(y_term(i+1)));

% We use the second equation of (E) to determine delta
exp3 = (rhat+(muhat-rhat)*pi0+beta-0.5*sigma^2*pi0^2*(1-rho))*rho;
exp4 = beta*gamma - lambda*((1-pi0*eta)^rho-1+rho*pi0*eta);
delta=exp3-exp4;
k = (1-rho)/(beta*(rho-gamma));

% We check that V_t + F(t,r,V_r,V_rr,I) <= 0 for the exact solution of our
% optimization problem, which is a necessary condition for the exact solution
     formula to
% be valid.
r0 = linspace(k,10,100);
pivec0=linspace(0,1,100);
[r,pivec]=meshgrid(r0,pivec0);
c = rho*(1-gamma)^(1-gamma)/(gamma*(delta+beta*gamma)*(rho-gamma)^(1-gamma));
exp3 = (1+beta.*r.*(1-pivec*eta)).^gamma - (1+beta*r).^gamma + pivec.*r.*eta
    .*gamma.*beta.*(1+beta*r).^(gamma-1);
exp3 = lambda*exp3;
exp2 = (rhat+(muhat-rhat).*pivec).*gamma.*beta.*r./(1+beta*r) + 0.5*sigma
    .^2.*pivec.^2.*((beta.*r)./(1+beta*r)).^2*gamma*(gamma-1) + exp3;
exp1 = 1/gamma - delta*c.*(1+beta*r).^gamma - beta.*gamma./(1+beta*r).*c.*(1+
    beta.*r).^gamma + c*(1+beta*r).^gamma.*exp2;
% maxValue = the value of r and pi that maximizes V_t + F(t,r,V_r,V_rr,I)
maxValue= max(max(exp1));

if maxValue>0
    display('not valid choice of parameters')
    delta=-100; rho=-100;
end
end
```

```
% boundary condition for r=0
function v=r0(t,k1,delta)
    v = exp(-delta*t).*k1;
end
```

```
% terminal condition
function v=tT(T,x,k,k1,k2,k3,delta,gamma,beta,rho)
    index=x<k;
    v=x*0;
    v(index) = k1 + k2*(x(index)/k).^rho;
    v(~index) = k3*(1+beta*x(~index)).^gamma/(1+beta*k)^gamma;
    v = exp(-delta*T)*v;
end
```

```
% calculates the integral part of the HJB equation for case A'
function jump = calcJump_caseB(pi,Rmax,V,xsi0)
global xsi;
xsi=xsi0; % size of the jump

Nr=length(pi)+1; % number of grid points minus 1
dr=Rmax/Nr; % delta r
r=linspace(dr,Rmax-dr,Nr-1)'; % vector containing interior values of r
sind0=1:(Nr-1);

A = zeros(length(r),Nr+1); % coefficients of each v_i in the calculation of
    the jump
dz = sqrt(dr);

Nmax = ceil(xsi/dr-1); % the number of terms in the sum approximating the
    integral

% calculating the double integral of k for all relevant n
kTilde = zeros(Nmax+1,1); % the double integral of k
n=1:(Nmax+1);
start=-n*dr;
start(end)=-xsi;
stop=-(n-1)*dr;
kTilde(n) = 0.5.*stop.^2 - 0.5.*start.^2 + xsi.*(stop-start);

% For each term of the sum approximating the integral, we calculate D_X^2v
    numerically by a three point method. The three points will be called 'r_m
    ', 'r_z' and 'r_p' below. Each of the three points 'r_x' (x=m,z,p) lie
    between two points r_i and r_{i+1} of our grid. The vectors 'wp1', 'wm1'
    and 'wz1' give us the weight of v(r_i) in the linear interpolation where
    we approximate v(r) by v(r_i) and v(r_{i+1}). The vectors 'lm', 'lz' and
    'lp' denote the index i for r_m, r_z and r_p, respectively.
wp1 = zeros(length(pi),Nmax+1);
wm1 = zeros(length(pi),Nmax+1);
wz1 = zeros(length(pi),Nmax+1);

n=1:(Nmax+1);
[zMesh,rMesh]=meshgrid(n,r);

% calculating 'wm' and 'lm'
[zMesh,piMesh]=meshgrid(-dr*(n-0.5)-dz,pi);
wm = eta(rMesh,zMesh,piMesh)/dr;
lm = floor(wm);
wm = wm - lm;
```

```
% calculating 'wp' and 'lp'
[zMesh,piMesh]=meshgrid(-dr*(n-0.5)+dz,pi);
wp = eta(rMesh,zMesh,piMesh)/dr;
lp = floor(wp);
wp = wp - lp;

% calculating 'wz' and 'lz'
[zMesh,piMesh]=meshgrid(-dr*(n-0.5),pi);
wz = eta(rMesh,zMesh,piMesh)/dr;
lz = floor(wz);
wz = wz - lz;

% removing contributions corresponding to points outside our grid
sind0Mesh = meshgrid(sind0,ones(Nmax+1,1))';
index = (sind0Mesh+lp+1<Nr & sind0Mesh+lm>1);
lm(~index)=0;
wm(~index)=0;
wm1(index)=1-wm(index);
lp(~index)=0;
wp(~index)=0;
wp1(index)=1-wp(index);
lz(~index)=0;
wz(~index)=0;
wz1(index)=1-wz(index);

kTildeMMesh=meshgrid(kTilde,ones(size(r)));
sind0Mesh=meshgrid(sind0,lm(1,:))';
length_r=length(r);

% adding contributions to the A-matrix from the points 'r_m'
sindNew1=zeros(size(lm));
sindNew2=zeros(size(lm));
sind1 = sind0Mesh + (lm+1);
sind2 = sind0Mesh + (lm);
for i=1:size(lm,1)
    sindNew1(i,:) = sind1(i,:)*length_r+i-1;
    sindNew2(i,:) = sind2(i,:)*length_r+i-1;
end
for n=1:(Nmax+1)
    A(sindNew1(:,n)+1) = A(sindNew1(:,n)+1)+wm(:,n).*kTildeMMesh(:,n)/dz^2;
    A(sindNew2(:,n)+1) = A(sindNew2(:,n)+1)+wm1(:,n).*kTildeMMesh(:,n)/dz^2;
end

% adding contributions to the A matrix from the points r_p
sind1 = sind0Mesh + (lp+1);
sind2 = sind0Mesh + (lp);
for i=1:size(lm,1)
    sindNew1(i,:) = sind1(i,:)*length_r+i-1;
    sindNew2(i,:) = sind2(i,:)*length_r+i-1;
end
for n=1:(Nmax+1)
    A(sindNew1(:,n)+1) = A(sindNew1(:,n)+1)+wp(:,n).*kTildeMMesh(:,n)/dz^2;
    A(sindNew2(:,n)+1) = A(sindNew2(:,n)+1)+wp1(:,n).*kTildeMMesh(:,n)/dz^2;
end

% adding contributions to the A matrix from the points r_z
sind1 = sind0Mesh + (lz+1);
sind2 = sind0Mesh + (lz);
for i=1:size(lm,1)
    sindNew1(i,:) = sind1(i,:)*length_r+i-1;
    sindNew2(i,:) = sind2(i,:)*length_r+i-1;
end
for n=1:(Nmax+1)
    A(sindNew1(:,n)+1) = A(sindNew1(:,n)+1)-2*wz(:,n).*kTildeMMesh(:,n)/dz^2;
```

```
    A(sindNew2(:,n)+1) = A(sindNew2(:,n)+1)-2*wz1(:,n).*kTildeMMesh(:,n)/dz
        ^2;
end

% jump = our numerical approximation to the integral for each r
jump=A*V;
bTilde = r.*pi.*(xsi-1+exp(-xsi));
jump = jump - bTilde.*(V(2:(end-1),:)-V(1:(end-2),:))/dr;
end
```